ESTIMATION OF EFFICIENT REGRESSION MODELS FOR APPLIED AGRICULTURAL ECONOMICS RESEARCH

by

Octavio A. Ramírez

Sukant K. Misra

and

Jeannie Nelson[1]

1 Associate Professors and M.S. student, respectively. Department of Agricultural and Applied Economics, Texas Tech University, Box 42132, Lubbock TX 79409-2132. Phone: (806) 742-2821, Fax (806) 742-1099, Email (corresponding author): Octavio.Ramirez@ttu.edu

2002 Annual Meeting of the American Agricultural Economics Association
Long Beach, CA, July 28-31, 2002

ABSTRACT

This paper proposes and explores the use of a partially adaptive estimation technique to improve the reliability of the inferences made from multiple regression models when the dependent variable is not normally distributed. The relevance of this technique for agricultural economics research is evaluated through Monte Carlo simulation and two mainstream applications: A time-series analysis of agricultural commodity prices and an empirical model of the West Texas cotton basis. It is concluded that, given non-normality, this technique can substantially reduce the magnitude of the standard errors of the slope parameter estimators in relation to OLS, GLS and other least squares based estimation procedures, in practice, allowing for more precise inferences about the existence, sign and magnitude of the effects of the independent variables on the dependent variable of interest. In addition, the technique produces confidence intervals for the dependent variable forecasts that are more efficient and consistent with the observed data.

Key Words: Efficient regression models, partially adaptive estimation, non-normality, skewness, heteroskedasticity, autocorrelation.

Multiple regression models are used in applied agricultural economics research with two main purposes: forecasting and making statistical inferences about the effect of exogenous variables on the dependent variable. Efficient estimation of the model coefficients is important in both cases. Slope parameter estimators with lower standard errors represent more precise measurements of the magnitude of the impacts of the exogenous variables on the dependent variable, and produce more reliable predictions.

Ordinary least squares (OLS) is widely used in empirical work because if the model's error term is normally, independently and identically distributed (n.i.i.d.), OLS yields the most efficient unbiased estimators for the model's coefficients, i.e. no other technique can produce unbiased slope parameter estimators with lower standard errors. Maximum likelihood (ML) based on the n.i.i.d. assumption is equivalent to OLS. Generalized least squares (GLS) can be used to improve estimating efficiency relative to OLS when the error term is heteroskedastic, autocorrelated or both. Even more efficient slope parameter estimators can be obtained in this case through ML, by specifying a likelihood function in which the error term is assumed normal, but not i.i.d. (Judge et al.). This is commonly known as "correcting" for heteroskedasticity or autocorrelation.

If the dependent variable and, thus, the error term is a continuous but not normally distributed variable, however, OLS (in the i.i.d.-error case) or normal-error ML (in the non-i.i.d.-error case) is not the most efficient way of estimating the slope parameters of a multiple regression model (Judge et al.). Since non-normal dependent variables are not uncommon in applied modeling work, Goldfeld and Quandt argue vehemently against the continued reliance on the assumption of error term normality for estimating regressions. Three approaches are currently available for estimating multiple regression models under non-normality: robust, partially adaptive, and adaptive estimation (McDonald and White).

Robust estimators are based on minimizing some function of a scale-adjusted error term that gives less weight to large error values. They can be asymptotically more efficient than OLS when the tails of the underlying error term distribution are "thicker" than the normal (McDonald and Newey). Least Absolute Deviation (LAD) is an example of a robust estimator that is asymptotically more efficient than OLS for

many thick-tailed error-term distributions. Other robust estimators include the M and the L-estimators (Bickel, 1976; Mosteller and Tukey, 1977; Bierens, 1981; and Dielmam and Pfaffenberger, 1982; Joiner and Hall, 1983). Fully adaptive estimators involve non-parametric estimates of the derivative of the log of the unknown density. They have the advantage of being asymptotically efficient, as the maximum likelihood estimator obtained using knowledge of the actual error term distribution. Given that the correct error term distribution is seldom known, fully adaptive estimation should be preferred when working with large samples, since it would produce the lowest possible standard errors. Examples of fully adaptive estimators include Hsieh and Manski's Adaptive Maximum Likelihood (AML) estimator, which is based on a normal kernel density; and Newey's generalized method of moments estimator. For more details about these techniques please see McDonald and White. The more favorable asymptotic properties of fully adaptive estimators, however, are meaningless when working with small samples.

Partially adaptive estimators are ML estimators based on specific families of error term distributions, in hopes that the assumed family is flexible enough to accommodate the shape of the true unknown distribution of the error. Partially adaptive estimators based on the t distribution (Prucha and Kelejian), the generalized t (McDonald and Newey), and the generalized beta distribution (McDonald and White) have been explored in the econometrics literature. Partially adaptive estimators are asymptotically efficient only if the true error term distribution belongs to the family of the assumed distribution.

However, McDonald and White show that partially adaptive estimators based on assuming a flexible family of densities that can accommodate a wide variety of distributional shapes can substantially outperform OLS and all available robust and fully adaptive estimators in small sample applications involving non-normal errors. In other words, when estimating a multiple regression model with a continuous but non-normally distributed dependent variable and a small sample size, a partially adaptive estimator based on a flexible distribution that can reasonably approximate the true underlying error term distribution would likely produce slope parameter estimators with the lowest possible standard errors, potentially much lower than OLS.

Intuitively, and improved modeling of the error term distribution, i.e. of the distribution of the deviations from the observations from the multiple regression hyper plane, allows for a higher degree of certainty about the location of the hyper plane, i.e. about the parameters determining that location. If, for example, the error term distribution is substantially right skewed, OLS (which implicitly assumes a normally distributed error) cannot account for the extreme positive deviations that characterize right-skewness. As a result, the location of the regression hyper plane would be less certain that when assuming an error term distribution that can account for these deviations.

The use of partially adaptive procedures for increasing slope parameter estimation efficiency through a more precise modeling of the error term distribution, however, has not been explored in the agricultural economics literature. Given the importance of obtaining more precise estimates of a model parameters for forecasting and statistical inferences, partially adaptive estimation techniques could be very useful in applied agricultural economics research. Applied researchers are increasingly better trained in basic econometric techniques, including the standard maximum likelihood estimation procedures required to implement partially adaptive estimation. This should facilitate the adoption of more efficient modeling techniques by applied researchers.

However, a practical shortcoming of the available partially adaptive estimators is that they were not designed to model heteroskedasticity or autocorrelation. Since, in addition to inefficient slope parameter estimators, unchecked heteroskedasticity or autocorrelation leads to biased and inconsistent standard error estimators, this limits the applicability of partially adaptive estimation when the error term is non-normally distributed and non-i.i.d. In other words, an applied researcher concerned about efficiency would have to choose between partially adaptive estimation or "correcting" for the non-i.i.d. error. In addition, available partially adaptive estimators cannot be straightforwardly used in multiple-equation (i.e. seemingly unrelated equation –SUR–) set up, which is a common procedure to increase estimation efficiency in relation to OLS.

This paper addresses the former issues by proposing and evaluating the theoretical and empirical performance of a partially adaptive estimator that can jointly model error term non-normality,

heteroskedasticity and/or autocorrelation in single and multiple equation specifications. The technique is based on assuming that an expanded form of the Johnson S_U family of distributions (Johnson, Kotz, and Balakrishnan) can approximate the true underlying error term distribution. The Johnson S_U family has been previously applied by Ramirez to simulate non-normally distributed yield and price distributions for agricultural risk analysis. Through Monte Carlo simulation assuming a variety of scenarios, it is shown that when the underlying error term is non-normally distributed and non-i.i.d., the proposed estimator can substantially increase slope parameter estimation efficiency in comparison to OLS, GLS (normal-error ML), and all other partially adaptive estimators available in the econometrics literature. The proposed technique is also validated and illustrated through two agricultural time series modeling applications.

The Estimator

The proposed partially adaptive estimator is obtained by assuming that the model's error term (**U**) follows the following expanded form of the Johnson S_U family of distributions:

(1) $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{U}$,

(2) $\mathbf{U} = \sigma\{\sinh(\Theta\mathbf{V})-F(\Theta,\mu)\}/\{\Theta G(\Theta,\mu)\}$, $\mathbf{V} \sim \mathbf{N}(\mu,1)$,

$F(\Theta,\mu) = E[\sinh(\Theta\mathbf{V})] = \exp(\Theta^2/2)\sinh(\Theta\mu)$, and

$G(\Theta,\mu) = [\{\exp(\Theta^2)-1\}\{\exp(\Theta^2)\cosh(-2\Theta\mu)+1\}/2\Theta^2]^{1/2}$;

where **Y** is an $n\times1$ vector of observations on the dependent variable; **X** is an $n\times k$ matrix of observations on k independent variables including an intercept; $\boldsymbol{\beta}$ is a $k\times1$ vector of intercept and slope coefficients; $-\infty<\Theta<\infty$, $-\infty<\mu<\infty$, and $\sigma>0$ are transformation parameters; and sinh(x) and cosh(x) are the hyperbolic sine and cosine functions, respectively. Using the results of Johnson, Kotz and Balakrishnan (pp. 34-38) it can be shown that in the model defined above:

(3) $E[\mathbf{U}] = 0$, $Var[\mathbf{U}] = \sigma^2$,

$Skew[\mathbf{U}] = E[\mathbf{U}^3] = S(\Theta,\mu) = -1/4w^{1/2}(w-1)^2[w\{w+2\}\sinh(3\Omega)+3\sinh(\Omega)]/G(\Theta,\mu)^{3/2}$,

$Kurt[\mathbf{U}] = E[\mathbf{U}^4] = K(\Theta,\mu) = \{1/8\{w-1\}^2[w^2\{w^4+2w^3+3w^2-3\}\cosh(4\Omega)+4w^2\{w+2\}\cosh(2\Omega)+3\{2w+1\}]/G(\Theta,\mu)^2\}-3$;

where $w = \exp(\Theta^2)$ and $\Omega = -\Theta\mu$. The equations in (3) indicate that, in this model specification, $\mathbf{X}\beta$ solely determines E[**Y**], σ^2 independently controls Var[**U**], and μ and Θ determine error term skewness and kurtosis. Thus, standard heteroskedastic specifications can be introduced by making σ^2 a function of the variables influencing Var[**U**], without affecting the E[**Y**] or the error term skewness or kurtosis. Evaluation of Skew[**U**] and Kurt[**U**] shows that if $\Theta \neq 0$ but $\mu = 0$ the distribution of **U** is kurtotic (i.e. "fat-tailed") but symmetric. The sign of Θ is irrelevant, but higher values of Θ cause increased kurtosis. If $\Theta \neq 0$ and $\mu > 0$, **U** has a kurtotic and right-skewed distribution, while $\mu < 0$ results in a kurtotic but left skewed distribution. Higher values of μ increase both skewness and kurtosis, but kurtosis can be scaled back by reducing $|\Theta|$.

In short, a wide variety of right and left skewness-kurtosis coefficient combinations can be obtained by altering the values of these two parameters. Also, if $\mu = 0$, $S(\Theta,\mu) = 0$, and the former becomes symmetric but kurtotic error term model. Further, as Θ goes to zero, **U** approaches $\sigma\mathbf{V}$, Var[**U**] approaches σ^2 and $K(\Theta,0)$ also becomes zero, indicating that the normal-error regression model is nested to this non-normal error model. As a result, in applied regression analysis, if the error term is normally distributed, both μ and Θ would approach zero and the proposed estimator for the slope parameter vector β would be the same as OLS. Also, the null hypothesis of error term normality (i.e. OLS) vs. the alternative of non-normality can be tested as Ho: $\Theta = 0$ and $\mu = 0$ vs. Ha: $\Theta \neq 0$ and $\mu \neq 0$. The null hypothesis of symmetric non-normality versus the alternative of asymmetric non-normality is Ho: $\Theta \neq 0$ and $\mu = 0$ vs. Ha: $\Theta \neq 0$ and $\mu \neq 0$.

To specify a non-normally distributed and autocorrelated error term model, consider a model with an $n \times 1$ error term vector **U**, which is normally distributed but not i.i.d. Following Judge, et al., let $\mathbf{\Phi} = \sigma^2\psi$ be the covariance matrix of the error term vector, **P** be an $n \times n$ matrix such as $\mathbf{P'P} = \psi^{-1}$, $\mathbf{Y^* = PY}$ (an $n \times 1$ vector), and $\mathbf{X^* = PX}$ (an $n \times k$ matrix), where **Y** and **X** are the vector and matrix of original dependent and independent variables. Given the choice of **P**, the transformed error term $\mathbf{U^* = PU = P(Y - X\beta)} = (\mathbf{PY - PX\beta}) = (\mathbf{Y^* - X^*B})$ is i.i.d. Under the assumption of normality, the log-likelihood function that has to be maximized in order to estimate a multiple regression model with non-i.i.d. errors then is:

(4) $$NLL_j = -(n/2)\ln(\sigma^2) - 0.5 \times \ln|\psi| - (\mathbf{U^{*\prime}U^*}/2\sigma^2).$$

A regression model that can accommodate non-normality (kurtosis and right or left skewness), and autocorrelation and/or heteroskedasticity, is finally obtained by applying the transformation in equation (2) to $\mathbf{U^*}$. When the error term vector $\mathbf{U} = (\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})$, is autocorrelated, a first transformation is used to obtain a non-autocorrelated error term $\mathbf{U^*} = \mathbf{PU} = (\mathbf{PY}-\mathbf{PX}\boldsymbol{\beta}) = (\mathbf{Y^*}-\mathbf{X^*}\boldsymbol{\beta})$. This is then transformed to a normal error term vector $\mathbf{V}$ through equation (2). If $\mathbf{U}$ is only heteroskedastic, $\mathbf{P}$ is an $n \times n$ identity matrix. The concentrated log-likelihood function to be maximized in order to estimate this model's parameters is:

$$\text{(5)} \qquad NNLL = -0.5 \times \ln|\boldsymbol{\psi}| + \sum_{i=1}^{n} \ln(G_i) - \sum_{i=1}^{n} 0.5 \times H_i^2 \}; \text{ where:}$$

$$G_i = G(\Theta,\mu)(1+R_i^2)^{-1/2}/\sigma; \qquad H_i = \{\sinh^{-1}(R_i)/\Theta\} - \mu; \qquad R_i = \{G(\Theta,\mu)\Theta U_i^*/\sigma\} + F(\Theta,\mu);$$

and $i=1,\ldots,n$ refers to the observations, $\sinh^{-1}(x) = \ln\{x+(1+x^2)^{1/2}\}$ is the inverse hyperbolic sine function, and $F(\Theta,\mu)$ is as given in equation (2). The first and second terms in equation (5) are the natural logs of the Jacobians of the first and second transformations, respectively. H_i is the inverse of the transformation to normality in equation (2). If the error term is believed to be autocorrelated, $\mathbf{P}$ and $|\boldsymbol{\psi}|$ must be specified to make equation (5) operational. Judge, et al. derives $\mathbf{P}$ and $|\boldsymbol{\psi}|$ for first- and higher-order autoregressive processes. As before, in this autocorrelated specification, $E[\mathbf{Y}] = \mathbf{X}\boldsymbol{\beta}$ and $Var[\mathbf{U}] = \sigma^2\boldsymbol{\psi}$. Thus, σ^2 can still be used to model systematic changes in the variance of $\mathbf{U}$ across the observations without affecting skewness or kurtosis.

The multiple equation (SUR) equivalent of equation (5) is obtained by applying a set of normality transformations [equation (2)] to a set of m "transformed" $n \times 1$ non-normal random errors, $\mathbf{U_j^*} = \mathbf{Y_j^*}-\mathbf{X_j^*}\boldsymbol{\beta_j}$ $(j=1,\ldots,m)$, where $\mathbf{Y_j^*}=\mathbf{P_jY_j}$, $\mathbf{X_j^*}=\mathbf{P_jX_j}$, and $\mathbf{Y_j}$ and $\mathbf{X_j}$ are the original vectors and matrices of dependent and independent variables, respectively. As in the single equation case, the transformed set of random vectors $\mathbf{V_j}$ follow a multivariate normal distribution with means μ_j $(j=1,\ldots,m)$ and covariance matrix $\boldsymbol{\Sigma}$. The non-diagonal elements of $\boldsymbol{\Sigma}$ (ρ_{jk}) account for the correlation between the error terms of the m equations. The concentrated log-likelihood function for this model is a straightforward multivariate extension of equation (5):

$$(6)\qquad MNNLL = \{ -(n/2)\times\ln|\boldsymbol{\Sigma}| - 0.5\times\sum_{j=1}^{m}[\ln(|\boldsymbol{\psi}_j|)] + \sum_{i=1}^{n}\sum_{j=1}^{m}[\ln(G_{ji})] - 0.5\times\sum_{i=1}^{n}\sum_{j=1}^{m}[\{\mathbf{H}_i*(\boldsymbol{\Sigma}^{-1})\}.*\mathbf{H}_i]\};$$

where $\boldsymbol{\Sigma}$ is an m×m matrix with unit diagonal elements and non-diagonal elements ρ_{jk}; G_{ji} is as defined in equation (5) if $\mathbf{Y_j}$ (and thus $\mathbf{Y_j^*}$) is not normally distributed or $G_{ji}=\sigma_j^{-1}$ if $\mathbf{Y_j}$ is normally distributed; and $\mathbf{H_i}$ is a 1xm row vector with elements H_{ji} (j=1,...,m) also defined in equation (5) if $\mathbf{Y_j}$ is not normally distributed and $H_{ji}=(Y_{ji}^*-X_{ji}^*\boldsymbol{\beta_j})/\sigma_j$ if $\mathbf{Y_j}$ is normally distributed. The operator $*$ indicates a matrix multiplication; and $.*$ indicates an element-by-element matrix multiplication.

The multivariate log-likelihood function {equation (6)} simply links m univariate functions [equation (5)] through the cross-error term correlation matrix $\boldsymbol{\Sigma}$. As in the normal error case, if some of the m dependent variables of interest are correlated to each other, using equation (6) to jointly estimate the $\boldsymbol{\beta_j}$ vectors should result in an improved efficiency in comparison with the case where they are estimated separately. Maximum likelihood estimation is conducted by finding the values of the parameters (Θ_j, μ_j, $\boldsymbol{\beta_j}$, and those in the $\boldsymbol{\Sigma}$, $\mathbf{P_j}$, and $\boldsymbol{\psi_j}$ matrices) that maximize the log-likelihood function [equation (6)]. This is achieved through numerical optimization procedures, such as the Newton-Raphson algorithm, which are available in most econometric software packages, including Gauss 386i. These pre-programmed procedures only require a few standard command lines and the log-likelihood function. In addition to parameter estimates, they provide standard errors based on a numerical estimate of the Hessian matrix of this function.

Monte Carlo Simulation Analysis

The sample design used by Hsieh and Manski, Newey, and McDonald and White was adopted for the Monte Carlo simulation to ensure comparability with previous results. For the first phase of the simulation, the regression model is given by:

$$(7)\qquad Y_{ji} = \beta_{j0} + \beta_{j1}X_{ji} + U_{ji} = -1 + X_{ji} + U_{ji}\ (j=1);$$

where the explanatory variable $X_{ji} = 1$ with a probability of 0.5 and $X_{ji} = 0$ with a probability of 0.5. X_{ji} is also assumed to be statistically independent of U_{ji}. Thus, each model can be interpreted as estimating a shift parameter that separates two identical distributions except for a location parameter. The specifications for

the error term distribution are taken from McDonald and White: (a) Normal {N(0,1)}; (b) Mixture of normals or variance-contaminated normal {0.9*N(0,1/9) +0.1*N(0,9)}; and (c) Lognormal. The distributions were re-scaled and shifted, when necessary, to be drawn from a parent population with unitary variance and zero mean. One thousand samples of size 50 were generated for each simulation experiment, using the same X's for each sample. For the second phase of the simulation, first order autocorrelation was induced by multiplying each of the simulated error terms by the inverse of the P matrix defined above, for two different ρ values of 0.5 and 0.8. Heteroskedasticity was induced by multiplying the errors by I+0.5, where I is a binomial index variable taking values of zero or one with equal probability. For modeling purposes I is assumed to be a known qualitative factor that shifts the variance of the dependent variable from 0.5 to 1.5. In both cases the error term distributions maintain a zero mean and unit variance.

Simulation Results and Estimator Performance

As in previous studies (McDonald and White; Newey; Hsieh and Manski) the root mean squared error (RMSE) of the slope estimators is used as the criteria for comparing the relative performance of different estimators. The results of McDonald and White for three underlying error term distributions: normal, kutotic-only and kurtotic and skewed (Table 1), are used as a basis for comparison. Under OLS, the RMSE of the slope estimators is always around 0.28. When the true error term is i.i.d. normal, the proposed estimator, yields maximum likelihood parameter estimates for Θ and μ that are zero or not statistically different from zero, and slope estimates that are identical to those from OLS. Most of the estimators explored in McDonald and White perform similarly or slightly worse than OLS, as expected, since OLS is the most efficient estimator under i.i.d. normal error-term conditions.

The variance-contaminated normal implies a symmetric unimodal but thick-tailed error-term distribution with a kurtosis coefficient of approximately 20 in this case. The proposed estimator performs relatively well in estimating the slope coefficient (β_2), with a RMSE of 0.115 (Table 2). Nine other estimating techniques provide comparable efficiency levels, but none produces a RMSE lower than 0.11 (Table 1).

A log-normal error term distribution is unimodal but not symmetric. It exhibits positive degrees of kurtosis and skewness. The RMSE of the proposed slope-parameter estimator is 0.050, approximately 20% of the RMSE of the OLS estimator. According to McDonald and White criteria, this is an "impressive" performance, only comparable to the EGB2's 0.05 RMSE. All other alternative estimation techniques produce RMSE's of 0.11 and above.

In summary, the estimator proposed in this study performs very favorably in comparison with the 16 estimating techniques evaluated by McDonald and White according to the standards applied by these authors and independently established in previous studies (Newey; Hsieh and Manski). It clearly excels when the underlying error-term distribution is asymmetric. This implies a potential for efficiency gains when using the proposed estimator under conditions of error-term non-normality and heteroskedasticity/autocorrelation, or in a disturbance related equations set up.

When the variance-contaminated error-terms are simulated to be heteroskedastic, the slope estimation efficiency (i.e. the RMSEs) of the EGB2 and of the proposed estimator are not substantially affected, and the modeling of heteroskedasticity with the proposed estimator, by letting $\sigma = \alpha 1 + \alpha 2 I$ (where I is the binomial index variable known to affect the error-term variance), only produces modest efficiency gains (Table 2). Under log-normal heteroskedastic errors, however, the EGB2 and the i.i.d form of the proposed estimator are substantially less efficient in comparison to the log-normal homoscedastic scenario. Their RMSEs increase from 0.054 and 0.050 to 0.117 and 0.104, respectively. In contrast, a RMSE of 0.202 is obtained when estimating a heteroskedastic but normally distributed error-term model specification by maximum likelihood. The modeling of heteroskedasticicy with the non-i.i.d. form of the proposed estimator recovers the lost estimation efficiency gains. The resulting RMSE of 0.052 is 25% of the RMSE obtained with the heteroskedastic normal model.

In contrast, failure to model autocorrelation (ρ=0.5 and ρ=0.8) reduces the slope estimation efficiency of all estimators. The RMSEs increase to 0.320 and 0.388, respectively, when OLS is used under a variance-contaminated autocorretaled error-term, and to 0.334 and 0.439 under a log-normal

autocorrelated error-term. The i.i.d. form of the proposed estimator limits those RMSE increases to 0.141 and 0.188 (variance-contaminated autocorrelated), and to 0.126 and 0.260 (log-normal autocorrelated). Comparable RMSEs are found under the EGB2.

In short, the efficiency gains in the estimation of the slope parameter resulting from the use of either the EGB2 or the i.i.d. from of the proposed estimator, in relation to OLS, are reduced to about 50% if the underlying error-term is non-normally distributed and autocorrelated. The relative reduction is even more severe in comparison to the standard maximum likelihood autocorrelated normal-error estimator and, in the 0.8 autocorrelation log-normal error term scenario, the later is actually more efficient (Table 2). This illustrates the importance of a partially adaptive estimator that can account for both error term non-normality and autocorrelation/ heteroskedasticity: The autocorrelated form of the proposed estimator fully recovers the efficiency losses, producing slope-estimator RMSEs that are similar to those obtained under the i.i.d. non-normal underlying error-term scenarios. These RMSEs are 50 to 75% lower than those obtained with the standard maximum likelihood autocorrelated normal estimator and 25 to 80% lower than those obtained under the EGB2 or the i.i.d. form of the proposed estimator (Table 2).

Agricultural Economics Applications

Simple Time Series Models of Agricultural Commodity Prices

An issue of interest for agricultural economists is whether real commodity prices have been declining through time and if price variability has changed over time making the production of a particular crop more risky. Both, the normal-error regression model and the proposed partially adaptive (i.e. non-normal error) regression model are used to analyze this issue in the case of annual (1950-1999) U.S. corn and soybean prices.

The two price series are adjusted for inflation to the year 2000 using the producer price index for all agricultural products (USDA/NASS, http://www.usda.gov /nass/, March 2001). Both series are stationary according to augmented Dickey-Fuller unit root tests. OLS models assuming second-degree polynomial time trends (i.e. $\mathbf{X}\boldsymbol{\beta} = \beta_0 + \beta_1 t + \beta_2 t^2$; t=1,…,50) are first estimated using Gauss 386i lreg procedure (Table 3). Durbin-Watson tests indicate first order positive autocorrelation in both cases. Therefore, standard tenth-order

autoregressive models are estimated using Gauss 386i autoreg procedure. Since the autocorrelation parameter estimates decrease in size and the second and higher-order autocorrelation coefficients are not statistically different from zero at the 10% level, it is concluded that a first order autoregressive [AR(1)] error term specification is sufficient to correct for autocorrelation.

The estimated AR(1) models are presented in Table 3. Note that, although the autocorrelation coefficients are statistically significant at the 1% level in both the corn and soybean price models, the standard error estimates for the estimators of the slope parameters β_{c1}, β_{c2}, β_{s1}, and β_{s2} are larger under the AR(1) models. This is a common occurrence in applied modeling work, not withstanding of the fact that, under non-independently distributed errors, the AR(1) slope parameter estimators are theoretically more efficient than OLS. The cause of this apparent contradiction is that, in the presence of autocorrelation, the OLS standard error estimators are biased, generally downwards, i.e. the OLS standard error estimates tend to underestimate the correct standard errors and, thus, are not reliable.

Theoretically, if corn and soybean prices are correlated through time, more efficient slope parameter estimators can be obtained by estimating the two models jointly, using a seemingly unrelated regression (SUR) procedure. This procedure is available in Gauss, SAS, and other statistical analysis software under the assumption of independently distributed errors, but not under autocorrelated errors. However, the log-likelihood function used to estimate such a model (available from the authors upon request) is a straightforward bivariate extension of equation (4). In the case of the soybean price model, the estimated standard errors for the estimators of the regression slope parameters are moderately lower under this SUR-AR(1) procedure, but they are slightly higher in the corn price model. The overall efficiency gains from SUR estimation are likely limited by the low corn-soybean price correlation coefficient estimate under the assumption of error term normality. The SUR-AR(1) model is then expanded to evaluate if the variances of the price distributions have been changing through time. The two most common time-dependent heteroskedastic specifications are evaluated:

(8) $\sigma_j = \sigma_{j0} + \sigma_{j1}t$, and $\sigma_j^2 = \sigma_{j0} + \sigma_{j1}t$, $t = 1,\ldots,50$, $j = c$ and s;

which make the standard deviation or the variance of the corn and soybean price distributions linear

functions of time, respectively. Likelihood ratio and t tests do not reject Ho: $\sigma_{j1} = 0$ at the 20% level of statistical significance, suggesting that the error terms are homoskedastic in both cases. Since β_{c1} and β_{c2} are not statistically significant at the 20% level, and second degree polynomial regression specifications are susceptible to multicollinearity, a final model [NSUR-AR(1)] is estimated without β_{c2} (Table 3), in which the linear time-trend parameter β_{c1} is statistically different from zero at the 5% level.

The NSUR-AR(1) model is the most statistically efficient model currently achievable by an applied researcher dealing with the simple issue of whether real commodity prices have been declining over time. However, the previously discussed Monte Carlo simulation results suggest that the non-normal regression modeling techniques described above offer the potential for increased efficiency, i.e. slope parameter estimators with lower standard errors. The NSUR-AR(1) model is expanded to allow for the possibility of error term skewness and kurtosis. This expanded model [NNSUR-AR(1)] is estimated by maximizing the concentrated log-likelihood function given in equation (6), setting m=2. This log-likelihood is a function of the same parameters as the NSUR-AR(1) model, plus the parameters accounting a potential non-normality in the error term distributions of the corn (Θ_c and μ_c) and soybean (Θ_s and μ_s) models.

The statistical significance of the non-normality parameters is verified through the most reliable likelihood ratio tests. The concentrated log-likelihood function for the NSUR-AR(1) model reaches a maximum value of 36.92, versus 52.55 in the case of the NNSUR-AR(1) model. Thus, the likelihood ratio test statistic for Ho: $\Theta_c=\mu_c=\Theta_s=\mu_s=0$ is $\chi^2_{(4)} = -2\times(36.92-52.55) = 31.26$ allows for rejection of Ho at the 1% level. The non-normal SUR-AR(1) regression model is statistically superior to the normal SUR-AR(1) model. If the NNSUR-AR(1) model is restricted to only allow for non-normality in the error term of the corn price regression, the maximum value of the log-likelihood function decreases from 52.55 to 42.89, resulting a $\chi^2_{(2)}$ test statistic of $-2\times(42.89-52.55) = 19.32$, and rejection of Ho: $\Theta_s=\mu_s=0$ at the 1% level. Alternatively, the $\chi^2_{(2)}$ test statistic for Ho: $\Theta_c=\mu_c=0$ is $-2\times(46.12-52.55) = 12.86$ rejects Ho at the 1% level as well. The previous results indicate that the error terms of both the corn and the soybean price models are non-normally distributed.

The powerful D'Agostino-Pearson omnibus (K2) normality test can be applied to the OLS residuals to confirm these results. This test is based on a standardized combination of the well-known skewness and kurtosis coefficients, which, on the one hand, allows for rejection of the null hypothesis of normality on either count (skewness or kurtosis) and, on the other, it avoids the "double jeopardy" of separate skewness and kurtosis tests. The K2 statistic is 11.93 for the corn price residuals and 8.26 for the soybean price residuals. Since under the null hypothesis of normality K2 is distributed as a $\chi^2_{(2)}$, normality is rejected at the 1% significance level in both cases. Non-normal error term models are strongly justified. However, the estimated value for the skewness parameter in the corn model (μ_c) is very low (-0.0539) and statistically insignificant at the 20% level. A likelihood ratio test ($\chi^2_{(1)} = -2\times(52.54-52.55) = 0.02$) confirms this result. Therefore, μ_c is set equal to zero in the final NNSUR-AR(1) model (Table 3).

As explained earlier, $\mu_c=0$ implies that the distribution of the error term in the corn model is kurtotic but not skewed, while the positive sign of the parameter estimate for μ_s indicates positive, i.e. right skewness in the error term distribution of the soybean model. The specific skewness and kurtosis coefficients, calculated using the formulas in equation (3), are 0 and 22.20 (corn), and 2.08 and 8.61 (soybeans), respectively. Right skewness, i.e. upward price spikes that are relatively more pronounced than lower price occurrences, is not a surprising finding, and can be visually perceived on a scatter plot of the soybean price data (Figure 1). The simulation results are empirically validated in this case: The estimated standard errors for the slope parameter estimators, are substantially lower under the non-normal model (0.39, 2.81, and 5.30 vs. 0.56, 3.83 and 7.50 under the normal), which means that the proposed partially adaptive estimator allows for much narrower (i.e. more precise) confidence intervals for the rates of decreases of corn and soybean prices through time.

Confidence intervals for the price occurrences involve the uncertainty in the estimation of all of the model parameters as well as the estimated error term distribution. These are obtained by applying the numerical technique of Krinsky and Rob. This technique is based on the asymptotic properties of the ML estimators, i.e. on the fact that they are normally distributed, consistent, and with known covariance matrix. It uses the same principle applied to construct confidence intervals for the true parameter values to

numerically simulate the probability distribution from which they could have been drawn. Specifically, let $\hat{\Gamma}_i$ be the k*x*1 vector of maximum likelihood estimators for Γ_i, the vector of true population parameters underlying the normal (i=N, k=10) or the non-normal (i=NN, k=14) regression model, and $\hat{CM}[\hat{\Gamma}_i]$ be the estimated covariance matrix for $\hat{\Gamma}_i$. Then, the joint probability distribution of $\hat{\Gamma}_i$ is simulated by:

$$S_i = \mathbf{Z}\ x\ \text{Chol}(\hat{CM}[\hat{\Gamma}_i]) + \tilde{\Gamma}_i, \tag{9}$$

where **Z** is an m*x*k matrix of independently distributed standard normal random variables, Chol(.) denotes the Cholesky decomposition, and $\tilde{\Gamma}_i$ the k*x*1 vector of parameter estimates obtained from $\hat{\Gamma}_i$. Equation (9) yields an m*x*k matrix of random variables with mean $\tilde{\Gamma}_i$ and covariance matrix $\hat{CM}[\hat{\Gamma}_i]$. Since, under a correct model specification, $\hat{\Gamma}_i$ is a consistent estimator for Γ_i and $\hat{CM}[\hat{\Gamma}_i]$ is a consistent estimator for $\hat{CM}[\hat{\Gamma}_i]$, S_i is a theoretically correct probabilistic statement about Γ_i. Thus, the boundaries of a $(1-\alpha)$% confidence interval for the expected price under the normal (non-normal) model at time period t can be obtained by extracting the m sets of simulated parameter values from S_N (S_{NN}), using them to obtain m "predicted" price values for time t, and finding the $(\alpha/2)\ x\ m^{th}$ and the $[(1-\alpha)+\alpha/2]\ x\ m^{th}$ largest of these m price values.

Confidence intervals for the actual price realizations require simulation of m error term draws as well. In the case of the normal regression model, these are obtained by extracting the m simulated values for the standard deviation parameter (σ_c or σ_s) from S_N and multiplying them by m independent draws from a standard normal random variable. In the case of the non-normal model, the m sets of simulated values for σ_c, μ_c and Θ_c (or σ_s, μ_s and Θ_s) have to be extracted from S_{NN} and coupled with m independent standard normal draws. Then, m non-normal error term values are simulated by applying equation (2). The final step in constructing the boundaries of a $(1-\alpha)$% confidence interval for the actual price observations is to add the m simulated error term values to the corresponding m price "predictions" and find the $(\alpha/2)\ x\ m^{th}$ and the [(1-

$\alpha)+\alpha/2]$ *x* m^{th} largest of the resulting m simulated price realization values. The former process is programmed in Gauss 386i with m=100,000 and repeated for t=0,…49 under the normal and non-normal corn and soybean price models. The resulting boundaries are joined to obtain (1-α)% confidence bands for the 50 individual price predictions and for the 50 individual price realizations observed from 1951 to 2000.

The 95% confidence bands for the soybean price predictions, i.e. for the expected soybean prices, and for the actual price occurrences under the normal and non-normal regression models are presented in Figures 1 and 2. The predictions, and the corresponding confidence bands, do not follow the typical curvilinear shape of a second-degree polynomial in this case because they are autocorrelated predictions (Judge, et al., p. 316). The 95% bands for the 50 soybean price predictions under the normal model show an average width of $0.714/bu versus $0.557/bu under the non-normal model, i.e. they are 28% wider, on average. The 95% confidence bands for the 50 corn price predictions are 35% wider under the normal model than under the non-normal model, on average. An evaluation of the 80% to 95% confidence bands for both price series finds similar (28% to 35%) average percentage width differences at all 16 certainty levels.

In short, the non-normal models produce more precise predictions for the corn and soybean prices to be expected at any given year. Realistic confidence bands for the actual price occurrences, however, are as important in applied research. Figures 1 and 2 also show the 95% confidence bands for the soybean price occurrences under the normal and non-normal models. In the normal case (Figure 1), the bands are wider than those constructed for the predictions, but they are still symmetric about the price predictions. This is due to the addition of the normally distributed error term. Note that two of the price realizations exceed the upper bound while none is located even close to the lower bound, versus the theoretically required number of 1.25.

Examination of the 80 to 95% confidence bands reveals that, from 83% to 95%, no observations are left below the lower bounds, and only two are beneath the 80% band versus the theoretically required number of five (Table 4). When all of the 16 confidence bands are considered, only five out of a theoretically required total of 50 observations are found below the lower bounds, while 58 are above the upper bounds (Table 4). The assumption of error term normality, which causes these confidence bands to be symmetric, appears

incompatible with the soybean price data. The confidence bands for the soybean price occurrences under the non-normal model (Figure 2) are markedly non-symmetric about the predictions, reflecting the kurtosis and right-skewness of the estimated error term distribution. The 95% band leaves one observation above and one below its boundaries (1.25 required). The, 90, 85 and 80% bands leave three and two, five and three, and six and four observations below and above their boundaries, respectively, vs. the theoretically required numbers of 2.5, 3.75 and five, respectively. When all of the 16 confidence bands are considered, 59 and 42 observations are found below and above their lower bounds, respectively, versus the theoretically required number of 50 (Table 4). Compatibility with the data is improved by lower bounds that are consistently higher than their normal counterparts, and thus closer to the low price occurrences, combined with upper bounds that become relatively higher than the normal bounds at reduced α levels (Table 2 and Figures 1 vs. 2). The non-normal confidence bands are also narrower than the normal bands, on average and up to the 93% confidence interval.

On average, the 16 non-normal confidence bands for the corn price occurrences also appear more compatible with the observed data than the bands under the normal-error model, leaving 52 and 46 observations below and above their lower bounds, respectively, vs. 28 and 45, respectively. They are also narrower than the normal bands on average and up to the 93% confidence interval (Table 4).

The fact that the non-normal models result in confidence bands that are more consistent with the observed data is important given the many empirical applications of confidence intervals. It also provides for an intuitive explanation of the more precise slope parameter and dependent variable predictions afforded by the non-normal models. That is, the degree of uncertainty about the location of the true regression line (or hyper plane in the multiple regression case) is reduced by an improved accounting of the phenomena causing the data deviations form the line (or hyper plane), i.e. by the improved modeling of the error term distribution.

An Empirical Model of the West Texas Cotton Basis

Another issue of importance to agricultural economists is the behavior of the basis in a futures market, and measuring the impact of market and policy factors on the level and variability of the basis (Seamon and Kahl). For the purposes of this study, the monthly (January 1980 to December 2000) West

Texas cotton basis is defined as the difference between the West Texas cash price and the U.S. futures cotton price for the September contract.

Several factors have been hypothesized to affect this basis, including the monthly Texas, U.S., and foreign production (TXP, USP, and FP) and beginning stocks (TXBS, USP, and FBS); monthly foreign mill use (FMU); the U.S. price of rayon (PR); the opportunity cost of storage (STRC), as measured by Seamon and Kahl; transportation costs measured by the monthly U.S. railroad index for farm products (RRI); seasonal effects measured by dummy variables representing the planting (SD=0 for March to July) versus the harvesting and marketing seasons (SD=1 for August to February); and agricultural policy effects measured by dummy variables for the pre 1985 farm bill years (PD=1 from 1980 to 1985 and zero otherwise), the 1985 farm bill period (PD1=1 from 1986 to 1995 and zero otherwise), and the post 1996 farm bill era (PD2=1 from 1996 to 2001 and zero otherwise).

All of the variables discussed above are stationary according to the augmented Dickey-Fuller unit root test, with the exception of the price of rayon. The first difference in the price of rayon (FDPR), which is stationary, is therefore used instead of PR. A Durbin-Watson test based on the OLS residuals reveals first-order positive autocorrelation. The autocorrelation function of the OLS residuals is steadily decreasing, while the partial autocorrelations become statistically insignificant at lag five. When a tenth-order autoregressive model is estimated using Gauss 386i autoreg procedure, the autocorrelation parameter estimates decrease in size and the fifth and higher-order coefficients are not statistically different from zero at the 10% level. Thus, it is concluded that a fourth-order autoregressive [AR(4)] error term specification is sufficient to correct for autocorrelation.

The standard maximum likelihood procedures outlined above [equation (4)] are then used to estimate an AR(4) model under the assumption of error term normality and heteroskedasticity [NHAR(4)], shifting the variance of the error term (σ^2) by σ^2_{SD}, σ^2_{PD1}, and σ^2_{PD2}, according to the seasonal and policy dummies (Table 5). As in previous econometric analysis of regional basis, most of the regression parameters are insignificant. With the exception of SD, the relatively large standard errors of the slope parameter estimators make it impossible to

ascertain the direction and magnitude of the impact of the proposed explanatory variables on the basis, with reasonable statistical certainty.

The NHAR(4) model is then extended to account for error term non-normality [NNHAR(4)]. In this case, Ramirez and Shonkwiler's proposed procedures are expanded by letting the first non-normality parameter (Θ) shift by Θ_{SD}, Θ_{PD1}, and Θ_{PD2}, according to the seasonal and policy dummies. Since Θ affects both Skew[**U**] and Kurt[**U**] [equation (3)], this model specification assumes an error term distribution that can have a different variance, skewness and kurtosis depending on the season and policy period. The slope parameter estimators under this NNHAR(4) model exhibit substantially lower standard errors than their normal-error counterparts (Table 5), empirically validating the previously discussed Monte Carlo simulation results again. Final [FNHAR(4) and FNNHAR(4)] models are obtained by excluding the transportation cost (RRI) and 1996-2001 policy period (PD2) variables from the regression equation, since they show the lowest levels of statistical significance in both of the initial models and their corresponding parameter estimates bear incorrect signs.

All parameter estimates in the final models have the signs that would be expected from theory (Nelson). The estimates for the standard errors of the slope parameter estimators are again all higher under the normal model. Eight of the eleven standard error estimates are over 70% higher; and six are more than twice as high as in the non-normal model. On average, the standard error estimates under the NHAR(4) model are 99.6% higher. As a result, in the NNHAR(4) model, six regression parameters are statistically significant at the 1% level, and two more at the 5% level, while only one regression parameter is statistically significant under the NHAR(4) model.

These remarkable estimation efficiency gains are related to the substantial 70.47 increase in the maximum log-likelihood function value obtained by the modeling of non-normality. To put this number into perspective, recall that an increase of just $\chi^2_{(4,0.01)} \div 2 = 6.64$ is required for the likelihood ratio test to reject the null hypothesis that all non-normality parameters are equal to zero and conclude that the NNHAR(4) model represents a statistically significant improvement over the NHAR(4) model. Such a large log-likelihood value increase is an unmistakable indication that the observed basis data (Figure 3) is much more likely to have been generated by the non-normal than by the normal-error model. The D'Agostino-Pearson normality test applied to the OLS

residuals rejects the null hypothesis of error term normality at the 1% level, supporting this conclusion.

The confidence bands for the basis predictions and occurrences are generated using the same numerical procedures described in the case of the corn and soybean price models. As before, few key summary statistics about the bands are more revealing than their graphical presentation. As a result of the larger standard errors, the 80 to 99% confidence bands for the predictions from the normal model are, on average, 63% wider than those from the non-normal model, i.e. the later produces substantially more precise predictions of the West Texas cotton basis. As in the corn and soybean price models, the NNHAR(4) confidence bands for the actual basis occurrences are more "efficient" and theoretically consistent than the bands from the normal model (Table 6). They are more efficient in the sense that they are substantially narrower (7.306 vs. 9.643 basis points), on average, and at all 20 α levels evaluated, although, as before, they approach the width of the normal bands the lowest α of 0.01. At the same time, the 80 to 99% non-normal confidence bands are more theoretically consistent since the numbers of observations found below and above their lower and upper bounds are much closer to the theoretically expected numbers (Table 6). The more reliable non-normal confidence bands of the West Texas cotton basis are clearly useful for applied research and decision making about futures contracting.

Another interesting and original byproduct of the NNHAR(4) model is the finding that, in addition to affecting the mean and variance, policy and seasonal variables can also shift the skewness and kurtosis of the error term distribution and, thus, of the conditional distribution of an economic time series. The estimated conditional distributions of the West Texas cotton basis for the two seasons under the three policy periods in the analysis are presented in Figures 4 and 5, assuming the average values for all other explanatory variables during the corresponding season and policy period.

Note the substantial kurtosis and right skewness of the distributions during the March to July planting season. In the current, post-1985 Farm Bill period, for example, the distribution shows a mean of –2.62 and a standard deviation of 2.83. It implies a 0.3% (5%) probability of a basis realization that is less (more) than one standard deviation from the mean, versus the 16% expected under a normal distribution. The conditional distributions of the West Texas cotton basis before and during the 1985 Farm Bill period had estimated means

of –5.77 and –1.46 points, and standard deviations of 3.55 and 3.07, respectively. They were somewhat less kurtotic and asymmetric, with implied probabilities of basis realizations below (above) one standard deviation from their means of 5.9% and 7.1% (9.1% and 10.3%), respectively. In short, the probability of a high positive basis during the planting season was much higher during the 1985 Farm Bill period, and it is almost negligible in the current policy period, as it could be inferred from the data (Figure 3).

The estimated conditional distributions of the West Texas cotton basis during the harvesting and marketing season (Figure 5) are nearly normal under all the three policy periods analyzed. They show lower means (–6.08 vs. –5.77, –2.43 vs. –1.46, and –3.74 vs. –2.62, respectively) and standard deviations (2.64 vs. 3.55, 2.13 vs. 3.07, and 1.59 vs. 2.83, respectively) than during the planting season. Because of the stronger non-normality in the August-to-February distributions, however, these reduced standard deviations do not imply a lower overall degree of dispersion in the distribution of the basis in this case (Figures 4 and 5).

This is empirically important since, under error term non-normality, normal-error models such as the NHAR(4), provide consistent estimates for the error term variance (note the similarity in the variance estimates from the FNHAR(4) and the FNNHAR(4) models in Table 5), which would be misinterpreted as appropriate measures of the degree of dispersion of the distribution. For example, from the relatively large negative value of the variance shifter for the harvesting and marketing season (σ^2_{SD}) in the FNHAR(4) model, one would falsely conclude that the distribution of the basis during this season exhibits a much lower overall degree of dispersion about its mean value.

Concluding Remarks

Agricultural economics research often involves the estimation of regression models with a limited amount of data, and more precise and realistic statistical inferences from these models are always useful. As illustrated above, the non-normal error multiple regression model evaluated in this study can provide for substantially more precise and realistic statistical inferences than the currently available estimation techniques that assume normality. Since many dependent variables of interest to agricultural economists, such as commodity prices, crop acreage, yields, product supply, profits, etc., are likely non-normally

distributed, this could become an important tool in agricultural economics research. The theory and examples above also suggest that the degree of improvement over the normal-error models depends on how much the true error term distribution deviates from normality and on how well the non-normal error term distribution on which the partially adaptive estimation procedure is based approximates the true data-generating distribution. This is why very flexible distributions, such as the expanded form of the Johnson S_U family, should be preferred for partially adaptive estimation. The applications above also illustrate that proposed technique could be useful if the estimated models will be used to simulate conditional probability distributions for the dependent variable, which are often used for economic risk analysis.

References

Bickel, P.J. (1976). Another look at robustness. *Scandinavian Journal of Statistics*, 3:145-168.

Bierens, H.J. (1981). *Robust Methods and Asymptotic Theory in Nonlinear Econometrics*, Springer Verlag, Berlin.

D'Agostino, R.B., A. Belanger, and R.B. D'Agostino Jr. A Suggestion for Using Powerful and Informative Tests of Normality. *The American Statistician*, 44(4, 1990):316-321.

Dielmam, T., and Pfaffenberger (1982). LAV (least absolute value) estimation in linear regression: a review. *TIMS Studies in the Management Sciences*, 19:31-52.

Goldfeld, S.M., and R.E. Quandt. Econometric Modeling with Non-Normal Disturbances. *Journal of Econometrics*, 17(1981):141-155.

Hsieh, D.A., and C.F. Manski. Monte Carlo evidence on adaptive maximum likelihood estimation of a regression. *Annals of Statistics*, 15(1987):541-551.

Johnson, N.L., S. Kotz, and N. Balakrishnan. 1994. *Continuous Univariate Distributions*. New York: Wiley & Sons.

Joiner, B.L., and D.L. Hall (1983). The ubiquitous role of f'/f in efficient estimation of location. *The American Statistician*, 37:128-133.

Judge, G.G., W.E. Griffiths, R. Carter Hill, H. Lutkepohl, and Tsoung-Chao Lee. The *Theory and*

Practice of Econometrics, John Wiley & Sons, Inc., 1985.

Krinsky, I., and A.L. Robb. On Approximating the Statistical Properties of Elasticities. *Review of Economics and Statistics*, 68(1986): 715-719.

McDonald, J.B., and S.B. White. A Comparison of some Robust, Adaptive, and Partially Adaptive Estimators of Regression Models. *Econometric Reviews*, 12(1, 1993):103-124.

McDonald, J.B., and W.K. Newey. Partially Adaptive Estimation of Regression Models Via the Generalized t Distribution. *Econometric Theory*, 4(1988):428-457.

Moss, C.B., and J.S. Shonkwiler. Estimating Yield Distributions Using a Stochastic Trend Model and Non-Normal Errors. *American J. Agricultural Economics*, 75(November 1993): 1056-62.

Mosteller, F., and J.W. Tukey (1977). *Data Analysis and Regression*, Addison-Wesley, Reading, Massachusetts.

Nelson, J. An Empirical Investigation of the Cotton Basis for the Southern High Plains of Texas. MS Thesis, Dept. of Agricultural and Applied Economics, Texas Tech University, 2001, p. 36-40.

Newey, W.K. (1988). Adaptive estimation of regression models via moment restrictions. *Journal of Econometrics*, 38:301-339.

Ramirez, O.A. Estimation and Use of a Multivariate Parametric Model for Simulating Heteroskedastic, Non-Normal Random Variables: The Case of Corn-Belt Corn, Soybeans and Wheat Yields. *American J. Agricultural Economics* 79(February 1997):191-205.

Seamon, V.F. and K.H. Kahl. An Analysis of Factors Affecting the Regional Cotton Basis. Paper presented at the *NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management*; Chicago, Illinois, April 17-18, 2000.

Zeckhauser, R., and M. Thompson. Linear Regression with Non-Normal Error Terms. *Review of Economics and Statistics*, 52(1970):280-286.

Table 1. Root Mean Squared Error (RMSE) of slope estimators (sample size = 50; number of samples = 500). Source: McDonald and White (1993).

	Underlying Error-Term				Underlying Error-Term		
Estimation Technique	Normal	Normal Mixture	Log-Normal	Estimation Technique	Normal	Normal Mixture	Log-Normal
OLS	0.28	0.28	0.28	Huber 1, c=1	0.29	0.14	0.16
LAD	0.35	0.13	0.17	Huber 1, c=1.5	0.28	0.16	0.18
BT	0.29	0.14	0.18	Huber 1, c=2	0.28	0.19	0.20
GT	0.30	0.12	0.12	Huber 2 c=1	0.56	0.12	0.12
T	0.28	0.11	0.12	Huber 2 c=1.5	0.41	0.11	0.15
BT, p≥1	0.29	0.13	0.17	Huber 2 c=2	0.32	0.13	0.16
GT, p≥1	0.30	0.12	0.12	Manski (AML)	0.28	0.12	0.13
EGB2(p=q)	0.28	0.12	0.15	Newey (j)	0.30	0.12	0.11
EGB2	**0.29**	**0.12**	**0.05**	**Proposed**	**0.28**	**0.11**	**0.05**

Notes: OLS is the Ordinary Least Squares estimator; LAD is the Least Absolute Deviations estimator (Gentle, 1997); BT is the power exponential or Box-Tiao estimator (Zeckhauser and Thompson, 1970), GT is a partially adaptive estimator based on the generalized t distribution (McDonald and Newey, 1984, 1988); t is a partially adaptive estimator based on the Student's t distribution; EGB2 is a partially adaptive estimator based on the exponential generalized beta distribution of the second kind; Huber 1 and 2 refer to the estimators proposed by Huber (1964) and Huber (1981); Manski (AML) is the adaptive maximum likelihood estimator advanced by Hsieh and Manski (1987) based on a normal kernel density; and Newey (j) is the generalized method of moments estimator with j moments used in estimation (Newey, 1988). For more details about the former estimation techniques please see McDonald and White (1993).

Table 2: RMSE of the 1000-sample slope-parameter estimates under eigth underlying error-term distributions and the normal and the proposed partially adaptive estimator alternatives.

	Normal Estimators			**EGB2**	**Proposed Estimator**		
Underlying Distribution	i.i.d.	Heter.	Autoc.	i.i.d.	i.i.d.	Heter.	Autoc.
i.i.d. Var. Cont. Normal	0.284	--	--	0.126	0.115	--	--
i.i.d. Log-Normal	0.291	--	--	0.054	0.050	--	--
Het. Var. Cont. Normal	0.293	0.212	--	0.125	0.112	0.107	--
Het. Log-Normal	0.281	0.202	--	0.117	0.104	0.052	--
0.5 Aut. Var. Cont. Normal	0.320	--	0.260	0.149	0.141	--	0.114
0.8 Aut. Var. Cont. Normal	0.388	--	0.208	0.195	0.188	--	0.098
0.5 Aut. Log-Normal	0.334	--	0.238	0.143	0.126	--	0.053
0.8 Aut. Log-Normal	0.439	--	0.219	0.299	0.260	--	0.051

Table 3. Parameter estimates, standard error estimates, and statistical significance of parameters of normal and non-normal error regression models for corn and soybean prices.

	OLS		**AR(1)**		**SUR-AR(1)**		**NSUR-AR(1)**		**NNSUR-AR(1)**	
	Par. Est.	**S.E. Est.**	**Par. Est.**	**S.E. Est.**	**Par. Est.**	**S.E. Est.**	**Par. Est.**	**S.E. Est.**	**Par. Est.**	**S.E. Est.**
θ_C	--	--	--	--	--	--	--	--	0.9446	0.2875**
μ_C	--	--	--	--	--	--	--	--	--	--
B_{C0}	3.0840	0.1423**	3.0860	0.2263**	3.0817**	0.2273	3.1351	0.1600**	3.1722	0.1065**
B_{C1}	-1.5611	1.3432ns	-1.4552	2.1317ns	-1.4237ns	2.1879	-2.1217	0.5563**	-2.3509	0.3929**
B_{C2}	-1.0001	2.6509ns	-1.3787	4.1937ns	-1.4244ns	4.3186	--	--	--	--
σ_C	0.3490	--	0.2811	--	0.2818**	0.0282	0.2821	0.0282**	0.2990	0.0804**
ρ_C	--	--	0.5460	0.1185**	0.5347**	0.1111	0.5353	0.1112**	0.5688	0.0987**
θ_S	--	--	--	--	--	--	--	--	0.5677	0.1600**
μ_S	--	--	--	--	--	--	--	--	15.7161	5.0376**
B_{S0}	5.2544	0.3246**	5.3045	0.4998**	5.2899	0.4291**	5.3262	0.4172**	5.3607	0.3347**
B_{S1}	11.8443	3.0636**	11.5373	4.7101**	11.6223	4.0533**	11.1429	3.8273**	10.1967	2.8094**
B_{S2}	-25.4146	6.0462**	-25.0758	9.2701**	-25.1588	7.9850**	-24.1720	7.4968**	-21.4259	5.3032**
σ_S	0.7960	--	0.6587	--	0.6645	0.0677**	0.6645	0.0677**	0.7369	0.1603**
ρ_S	--	--	0.5132	0.1214**	0.4066	0.1227**	0.4073	0.1232**	0.4484	0.0763**
ρ_{CS}	--	--	--	--	0.3598	0.1310**	0.3597	0.1311**	0.4644	0.1158**
			MVCLF	**33.86**	**MVCLF**	**36.98**	**MVCLF**	**36.92**	**MVCLF**	**52.54**
R^2	**R^2_C=0.44**	**R^2_S=0.28**	**R^2_C=0.61**	**R^2_S=0.48**	**R^2_C=0.61**	**R^2_S=0.47**	**R^2_C=0.61**	**R^2_S=0.47**	**R^2_C=0.61**	**R^2_S=0.47**

Notes: MVCLF stands for the maximum value reached by the concentrated log-likelihood function. Par. Est. and S.E. Est. refer to the parameter and standard error estimates, respectively. The parameter and standard error estimates corresponding to B_{C1} and B_{S1}, and to B_{C2} and B_{S2} have been divided by 100 and 10000, respectively. * and ** denote statistical significance and the 90 and 95% level, respectively, according to two-tailed t tests. The R^2's are calculated by dividing the regression sums of squares (based on the autocorrelated {AR(1)} predictions) by the total sums of squares, i.e. it are the square of the correlation coefficients between the AR(1) predictions and the observed corn (c) and soybean (s) prices.

Table 4*. Select statistics about the 80 to 95% confidence bands for the price occurrences under the normal and non-normal corn and soybean price models.

		Normal Corn Price Model					Non-Normal Corn Price Model				
%CI	**TR#**	**#ULB**	**#OUB**	**ALB**	**AUB**	**AWCB**	**#ULB**	**#OUB**	**ALB**	**AUB**	**AWCB**
80%	5.00	4	3	2.249	2.993	0.744	7	3	2.315	2.910	0.595
81%	4.75	3	3	2.240	3.001	0.761	7	3	2.305	2.920	0.615
82%	4.50	3	3	2.231	3.010	0.779	7	3	2.294	2.931	0.636
83%	4.25	2	3	2.222	3.019	0.797	7	3	2.283	2.942	0.659
84%	4.00	2	3	2.213	3.029	0.817	5	3	2.270	2.955	0.684
85%	3.75	2	3	2.202	3.039	0.837	4	3	2.257	2.968	0.711
86%	3.50	2	3	2.191	3.050	0.859	4	3	2.243	2.982	0.740
87%	3.25	2	3	2.180	3.061	0.881	2	3	2.227	2.998	0.771
88%	3.00	1	3	2.168	3.074	0.905	2	3	2.210	3.015	0.806
89%	2.75	1	3	2.155	3.087	0.931	2	3	2.191	3.035	0.844
90%	2.50	1	3	2.141	3.100	0.959	1	3	2.169	3.056	0.887
91%	2.25	1	3	2.126	3.116	0.989	1	3	2.145	3.080	0.935
92%	2.00	1	3	2.109	3.132	1.023	1	3	2.117	3.108	0.991
93%	1.75	1	2	2.091	3.151	1.060	1	3	2.085	3.141	1.056
94%	1.50	1	2	2.070	3.171	1.101	1	2	2.046	3.180	1.134
95%	1.25	1	2	2.047	3.195	1.149	0	2	1.998	3.228	1.230
Sum/Ave	**50**	**28**	**45**	**2.165**	**3.077**	**0.912**	**52**	**46**	**2.197**	**3.028**	**0.831**
		Normal Soybean Price Model					**Non-Normal Soybean Price Model**				
%CI	**#TR**	**#ULB**	**#OUB**	**ALB**	**AUB**	**AWCB**	**#ULB**	**#OUB**	**ALB**	**AUB**	**AWCB**
80%	5.00	2	5	5.220	6.989	1.768	6	4	5.410	7.001	1.592
81%	4.75	2	5	5.200	7.009	1.809	5	4	5.401	7.035	1.634
82%	4.50	1	5	5.179	7.030	1.851	5	4	5.391	7.071	1.680
83%	4.25	0	5	5.157	7.052	1.895	5	4	5.382	7.110	1.728
84%	4.00	0	5	5.134	7.075	1.941	5	4	5.372	7.151	1.779
85%	3.75	0	5	5.109	7.099	1.990	5	3	5.362	7.195	1.833
86%	3.50	0	4	5.083	7.125	2.041	5	3	5.351	7.243	1.891
87%	3.25	0	4	5.057	7.152	2.095	5	2	5.340	7.295	1.955
88%	3.00	0	3	5.028	7.181	2.153	4	2	5.328	7.352	2.024
89%	2.75	0	3	4.997	7.212	2.214	3	2	5.316	7.414	2.098
90%	2.50	0	3	4.964	7.245	2.280	3	2	5.303	7.483	2.180
91%	2.25	0	3	4.929	7.280	2.352	3	2	5.289	7.561	2.272
92%	2.00	0	2	4.889	7.320	2.432	2	2	5.273	7.650	2.377
93%	1.75	0	2	4.844	7.364	2.520	1	2	5.256	7.752	2.496
94%	1.50	0	2	4.795	7.413	2.618	1	1	5.237	7.875	2.637
95%	1.25	0	2	4.738	7.470	2.732	1	1	5.216	8.021	2.805
Sum/Ave	**50**	**5**	**58**	**5.020**	**7.189**	**2.168**	**59**	**42**	**5.327**	**7.388**	**2.061**

*See notes below Table 5 for the definitions of the labels and summary statistics presented in Table 2.

Table 5. Parameter estimates, standard error estimates, and statistical significance of parameters of normal and non-normal error regression models for the West Texas cotton basis.

		NHAR(4)		**NNHAR(4)**		**FNHAR(4)**		**FNNHAR(4)**	
Par.	**Var.**	**P.E.**	**S.E.**	**P.E.**	**S.E.**	**P.E.**	**S.E.**	**P.E.**	**S.E.**
ρ_1	--	0.9807	0.0687[1]	0.96261	0.02720[1]	0.98904	0.06593[1]	0.95853	0.02773[1]
ρ_2	--	-0.2385	0.0965[2]	-0.13640	0.04190[1]	-0.24666	0.09358[1]	-0.13685	0.04244[1]
ρ_3	--	0.1319	0.0975	0.13418	0.04518[1]	0.13667	0.09606	0.13438	0.04602[1]
ρ_4	--	-0.1073	0.0612[3]	-0.15764	0.03732[1]	-0.11249	0.06017[3]	-0.15605	0.03682[1]
B_0	--	-14.5204	12.2673	-13.7478	8.60659	-18.26522	7.84887[2]	-11.83432	6.66378[3]
B_1	TXP	0.1930	0.5159	-0.11620	0.13143	0.18625	0.46566	-0.12095	0.12967
B_2	USP	0.3022	0.2476	0.31761	0.14303[2]	0.33011	0.23807	0.34039	0.08453[1]
B_3	FP	0.0913	0.1452	0.23745	0.08869[1]	0.10201	0.13689	0.24439	0.06782[1]
B_4	TXBS	-0.3323	0.3043	-0.47715	0.25815[3]	-0.32229	0.29186	-0.48446	0.24536[2]
B_5	USBS	0.1800	0.3508	0.52388	0.16576[1]	0.18375	0.34317	0.55098	0.15591[1]
B_6	FBS	0.0162	0.1053	0.16266	0.06824[2]	0.03877	0.10515	0.17332	0.04436[1]
B_7	FMU	-0.0361	0.3599	-0.26306	0.11320[2]	0.02261	0.17314	-0.26656	0.10223[1]
B_8	RRI	0.0146	0.1479	0.03185	0.08302	--	--		
B_9	STRC	-1.4566	1.2934	-1.15475	1.26164	-1.35748	1.25214	-1.13765	1.23339
B_{10}	FDPR	1.0527	0.6739	0.43163	0.47683	1.03744	0.67141	0.35747	0.38805
B_{11}	SD	-1.0991	0.4225[1]	-0.69499	0.23655[1]	-1.06572	0.41410[2]	-0.63615	0.18115[1]
B_{12}	PD1	2.1954	2.3614	1.85166	1.84490	1.29106	1.15962	2.38576	1.11527[2]
B_{13}	PD2	1.1976	2.6599	-0.85989	2.30983	--	--		
σ^2	--	16.0036	2.8967[1]	11.35527	4.13477[1]	16.13068	3.00359[1]	10.95924	3.63400[1]
σ^2_{SD}	SD	-11.3179	2.8122[1]	-5.82120	3.87761	-11.50518	2.77815[1]	-5.42047	3.34320[3]
σ^2_{PD1}	PD1	-0.3919	1.3692	-1.76188	1.42804	-0.38268	1.54083	-1.80659	1.40307
σ^2_{PD2}	PD2	-3.4360	1.0000[1]	-3.91090	1.35165[1]	-3.35245	1.02157[1]	-3.89927	1.33951[1]
Θ	--	--	--	0.95789	0.22514[1]	--	--	0.93976	0.20077[1]
μ	--	--	--	0.69869	0.25017[1]	--	--	0.72557	0.26177[1]
Θ_{SD}	SD	--	--	-1.21066	0.30811[1]	--	--	-1.19269	0.29155[1]
Θ_{PD1}	PD1	--	--	0.02153	0.19116	--	--		
Θ_{PD2}	PD2	--	--	0.57618	0.19320[1]	--	--	0.57207	0.17190[1]
MVLLF		**-352.820**		**-282.353**		**-352.964**		**-282.448**	
R^2		**0.77**		**0.77**		**0.77**		**0.77**	

Notes: NHAR(4), NNHAR(4), FNHAR(4) and FNNHAR(4) refer to the initial normal and non-normal and to the final normal and non-normal heteroskedastic fourth-order autoregressive models; the parameters (Par.) and variables (Var.) are as defined in the text; P.E. and S.E. refers to the parameter and standard error estimates; [1], [2], and [3] denote statistical significance at the 1%, 5% and 10% level, respectively, according to two-tailed t-tests; MVLLF refers to the maximum value reached by the model's concentrated log-likelihood function; and the R^2 is calculated by dividing the regression sum of squares (based on the autocorrelated {AR(4)} predictions) by the total sum of squares, i.e. it is the square of the correlation coefficient between the AR(4) predictions and the observed basis values.

Table 6. Select statistics about the 80 to 99% confidence bands for the basis occurrences under the final normal {FNHAR(4)} and non-normal {FNNHAR(4)} West Texas cotton basis models.

		FNHAR(4) MODEL					FNNHAR(4) MODEL				
%CI	**TR#**	**#ULB**	**#OUB**	**ALB**	**AUB**	**AWCB**	**#ULB**	**#OUB**	**ALB**	**AUB**	**AWCB**
80%	25.20	11	14	-6.483	0.697	7.180	21	27	-5.241	-0.456	4.785
81%	23.94	11	14	-6.566	0.779	7.345	19	24	-5.290	-0.370	4.920
82%	22.68	10	13	-6.652	0.865	7.516	19	22	-5.342	-0.278	5.064
83%	21.42	9	11	-6.741	0.954	7.695	19	21	-5.396	-0.181	5.215
84%	20.16	9	11	-6.835	1.048	7.883	18	20	-5.453	-0.076	5.377
85%	18.90	9	11	-6.934	1.146	8.081	17	20	-5.512	0.035	5.547
86%	17.64	9	11	-7.038	1.250	8.289	17	18	-5.577	0.156	5.733
87%	16.38	9	11	-7.149	1.361	8.509	16	15	-5.645	0.287	5.932
88%	15.12	7	8	-7.267	1.477	8.744	16	15	-5.719	0.431	6.150
89%	13.86	7	8	-7.392	1.601	8.994	16	14	-5.799	0.589	6.389
90%	12.60	5	7	-7.525	1.736	9.261	15	12	-5.886	0.764	6.650
91%	11.34	3	7	-7.672	1.882	9.554	14	12	-5.981	0.960	6.941
92%	10.08	3	7	-7.833	2.041	9.874	11	10	-6.089	1.185	7.274
93%	8.82	3	7	-8.012	2.219	10.230	7	9	-6.210	1.446	7.656
94%	7.56	3	5	-8.213	2.418	10.631	5	8	-6.350	1.753	8.104
95%	6.30	3	5	-8.447	2.649	11.096	5	7	-6.518	2.132	8.649
96%	5.04	3	4	-8.722	2.923	11.644	5	4	-6.725	2.613	9.338
97%	3.78	2	4	-9.065	3.263	12.328	3	4	-6.997	3.274	10.271
98%	2.52	2	1	-9.531	3.724	13.255	2	2	-7.399	4.288	11.687
99%	1.26	1	1	-10.28	4.473	14.755	1	1	-8.143	6.291	14.434
Sum/ Ave	**265**	**119**	**160**	**-7.718**	**1.925**	**9.643**	**246**	**265**	**-6.064**	**1.242**	**7.306**

Notes: #TR refers to the number of observations that would be theoretically required to be below and above the boundaries of the confidence band; #ULB and #OUB are the actual numbers found under and over the lower and upper bounds, respectively; ALB, AUB, and AWCB stand for the average of the n=50 lower and upper bounds and widths of the 50 confidence intervals comprising each of the 16 bands, respectively; and Sum/Ave. refer to the sums (for #TR, #ULB, and #OUB) or averages (for ABL, AUB, and AWCB) across the 16 confidence bands.

Figure 1: 95% Confidence Bands for the Soybean Price Predictions and Occurrences vs. Data under the Normal Regression Model

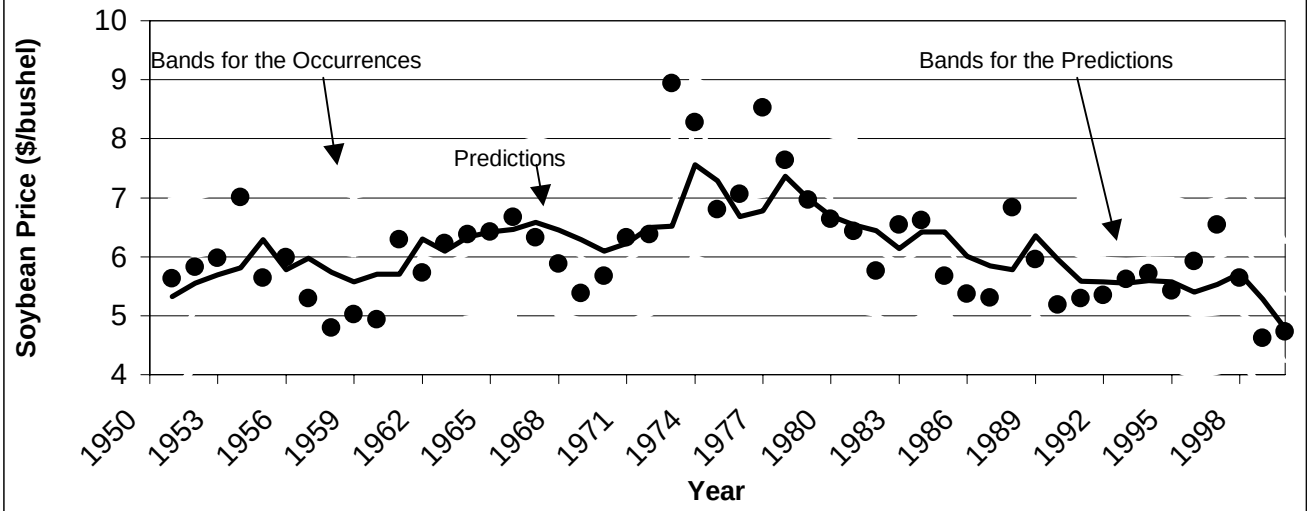
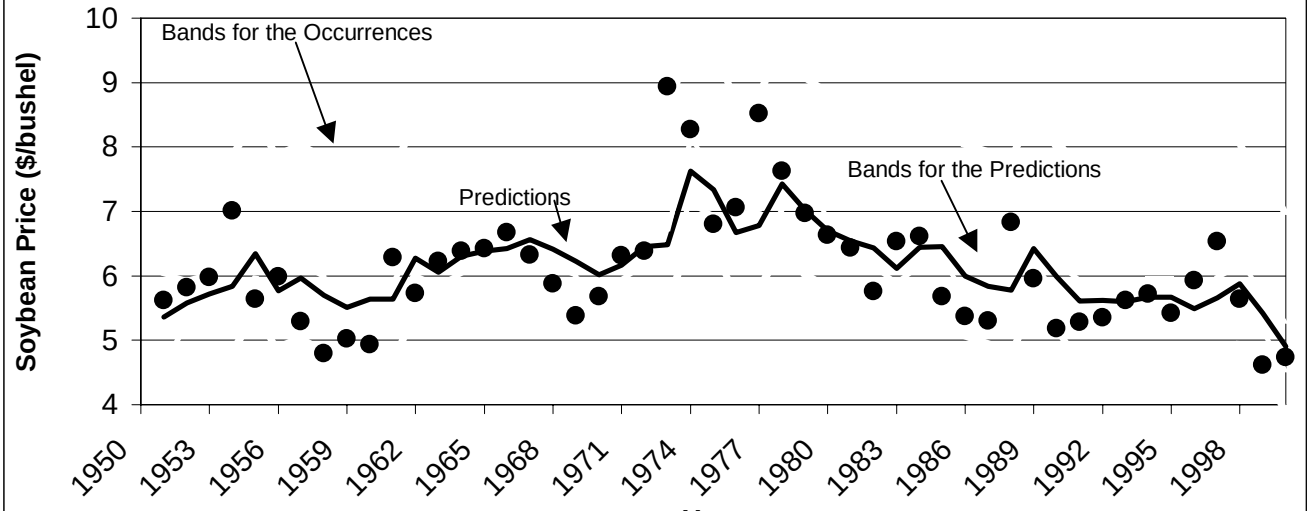


Figure 2: 95% Confidence Bands for the Soybean Price Predictions and Occurrences vs. Data under the Non-Normal Regression Model



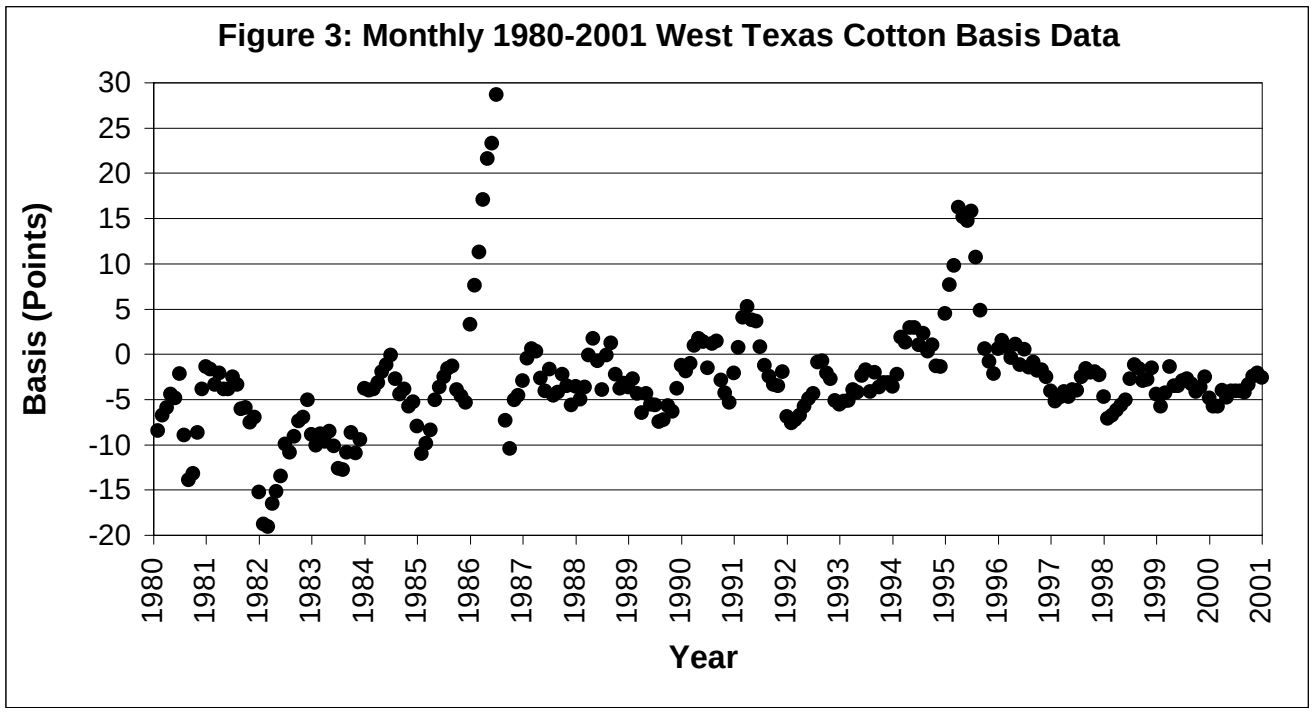


Figure 3: Monthly 1980-2001 West Texas Cotton Basis Data
Basis (Points)
30
25
20
15
10
5
0
-5
-10
-15
-20
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
Year

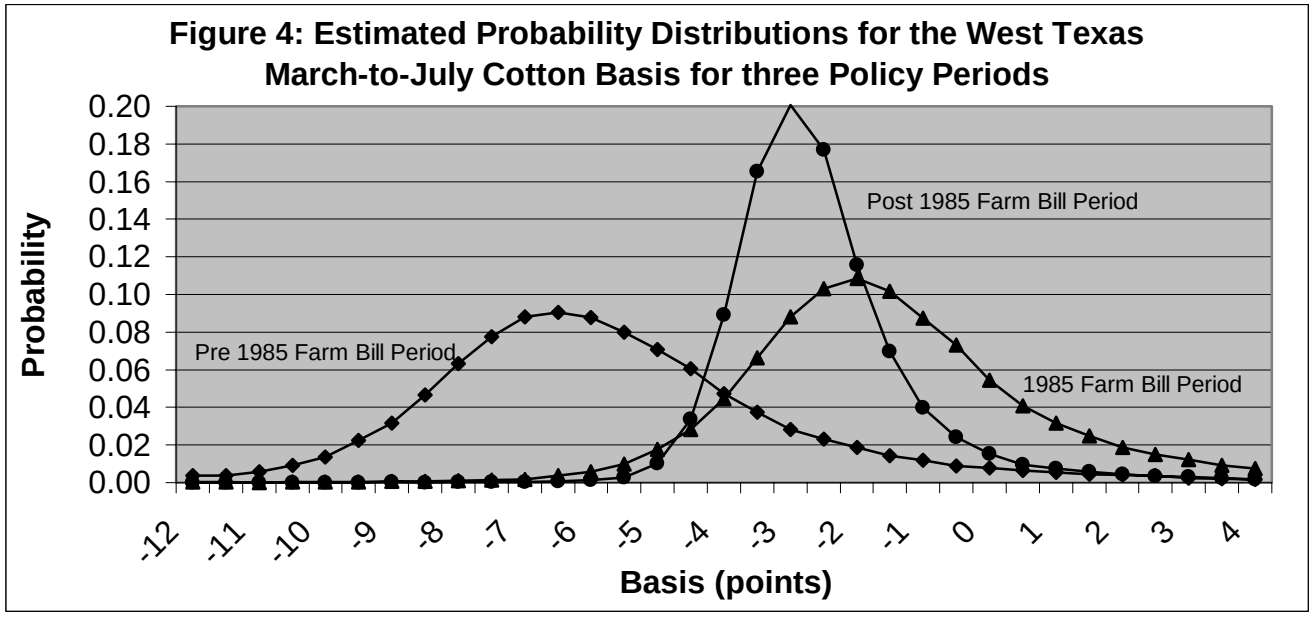


Figure 4: Estimated Probability Distributions for the West Texas March-to-July Cotton Basis for three Policy Periods
Probability
0.20
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
Post 1985 Farm Bill Period
Pre 1985 Farm Bill Period
1985 Farm Bill Period
-12
-11
-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
Basis (points)

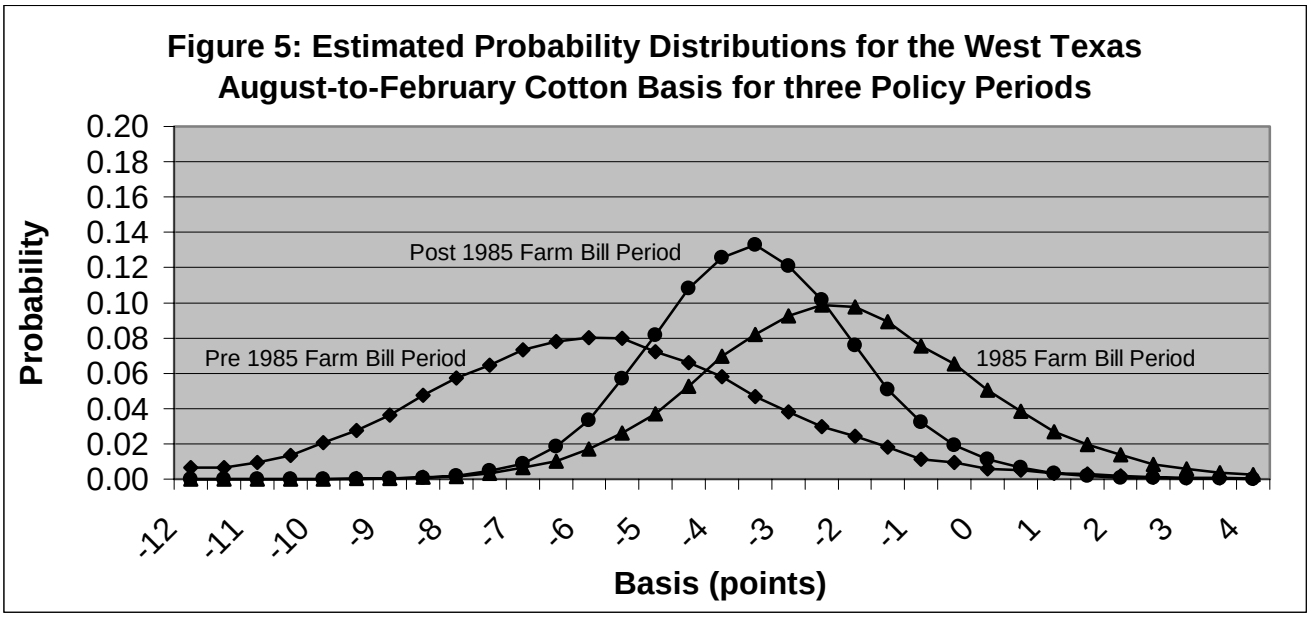


Figure 5: Estimated Probability Distributions for the West Texas August-to-February Cotton Basis for three Policy Periods
Probability
0.20
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
Post 1985 Farm Bill Period
Pre 1985 Farm Bill Period
1985 Farm Bill Period
-12
-11
-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
Basis (points)