WHAT CAN WE INFER ABOUT FARM-LEVEL CROP YIELD PDF's FROM COUNTY-LEVEL PDF's?

By

Zhiying Xu

A PLAN B PAPER

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Master of Science

Department of Agricultural Economics

2004

ABSTRACT

WHAT CAN WE INFER ABOUT FARM-LEVEL CROP YIELD PDF's FROM COUNTY-LEVEL PDF's?

By

Zhiying Xu

Accurate estimates of farm-level crop yield probability density functions (PDF's) are crucial for studying various crop insurance programs and production under risk and uncertainty. Unfortunately, farm-level crop yield PDF's are difficult to estimate due to the lack of sufficient farm yield data. County yield data cover much longer time periods than farm yield data, but using county yield distributions to conjecture about farm yield distributions is dangerous. The theoretical reason is that county yield is the *average* of correlated farm yields, for which there is no recognizable probability density function (PDF). This paper investigates the relationship between farm and county yield distributions using both statistical theory and the Monte-Carlo simulation method. Results show that under suitable farm yield correlation and density structures, the shape of yield distribution at the farm level is similar to that at the county level. A method is then developed for estimating and simulating farm yield distributions based on county yield PDF estimates and information contained in farm yield data.

Six candidate yield models: normal, beta, Weibull, inverse hyperbolic sine transformation, a mixture of normals, and kernel density estimators are applied to Branch County corn yields after detrending nonstationary yield data. Goodness-of-fit results for normal, beta and Weibull distributions show that Weibull best fits county yields. The method for simulating farm yields is illustrated using kernel density estimates.

ACKNOWLEDGEMENTS

I would like to acknowledge my major professor, Dr. Black for offering me guidance, assistance, and financial support through these years. I owe special debts to him for accepting and nurturing a fledgling research assistant. My sincere appreciation is also extended to my committee members, Dr. Steven Hanson and Dr. Jack Meyer for their helpful suggestions. I am deeply thankful to Professor Crawford for giving me the opportunity to study at such a stimulating place. Behind all my successful endeavors is the love, support and encouragement of my parents to whom I am always deeply grateful. Friends, colleagues, faculty and staff in the Department of Agricultural Economics make Michigan State University a happy place to study and do research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Producers of field crops are confronted with relatively high production risk compared to industrial manufacturers. Risk and uncertainty in crop yields stems from several sources. First, production of most crops often takes several months and yield is highly sensitive to many uncontrollable factors such as weather, pests and diseases. Second, varying crop management practices that can be controlled by farmers – adoption of new production techniques, input (for example, fertilizer) application level, timing of input application, and choice of varieties – is likely to result in high yield volatility. In addition, human and asset risks such as illness or death of a farm operator, loss or damage to the farm machinery and livestock may have significant impacts on crop yields.

Yield risk for a given crop can differ systematically over space due to changing agroecological conditions, mainly climate and soil type. Studies by Economic Research Service (ERS), USDA found that corn yield volatility[1] tends to be lowest in irrigated areas (such as Nebraska) and in the central Corn Belt (Iowa and Illinois) where climate and soils are ideal for corn production. It is typically higher outside the Corn Belt and in areas where corn acreage tends to be low.

A wide array of risk management tools and strategies are available for managing farm income risk which stems from yield and price risks, such as crop insurance, forward

[1] It is often measured by coefficient of variation (CV) indicator.

contracts, futures, options and crop diversification. Accurate characterizations of farm-level yield distributions, especially their lower tails, are important to many parties including farmers, insurance companies, lenders and the federal government: they are necessary for farmers to make sensible risk-management decisions, for insurance companies to precisely rate insurance premiums, and for lenders and the federal government to devise and provide farm risk management products.

Modeling and estimation of farm-level crop yield distributions is difficult for several reasons. First, historical farm yield data are available for at most 20 years and generally much less (Ker and Coble), which makes it difficult to estimate yield probability density function (PDF). Most studies resorted to county level or higher levels of aggregate time-series data that cover longer time periods, which may lead to improper representation of farm-level yield distributions. Yield volatility is likely to be lower at the county, district, state and national levels than at the individual farm level due to the averaging effect over the region of aggregation. Farm-level yield risk can be seriously underestimated with county-level or higher levels of data. Second, time-series crop yield data are usually found to be nonstationary, i.e., they are not the outcome of the same data generating process. Potential upward trend and increase in yield variance over time further complicates the estimation of yield PDF. Third, it has been recognized that crop yields may not be normally distributed in the relevant production range. However, there is little theoretical guidance regarding the most appropriate representation for the shape of the crop yield distribution.

The overall purpose of this paper is to investigate and simulate crop yield distributions at the farm level, expanding the existing literature by suggesting methodological improvements in assessing and simulating farm-level yield distributions when farm-level yield data are not sufficient. Specifically, the objectives of this paper are to: (i) apply multiple candidate yield models to assess yield distributions and compare these models; (ii) investigate the relationship between farm-level and county-level yield distributions using both statistical theory and Monte-Carlo simulation analysis; (iii) develop a new method for simulating farm-level yield distributions.

This paper is organized into eight chapters. Chapter 2 is literature review. Six candidate models are discussed in Chapter 3: normal, beta, Weibull, inverse hyperbolic sine transformation, a mixture of two normals, and nonparametric kernel estimators. Chapter 4 deals with detrending time-series yield data. The relationship between county-level and farm-level yield distributions is investigated in Chapter 5. Chapter 6 is a case study applying multiple candidate models. A new method for simulating farm-level yield distributions is developed in Chapter 7. Chapter 8 concludes with summary of the major results.

CHAPTER 2

LITERATURE REVIEW

The issue of modeling crop yield distributions has received attention in the agricultural economics literature since the 1960s. Day (1965) estimated yield distributions using Mississippi experimental data with seven nitrogen levels for cotton, corn and oats. He argued that crop yield distributions are nonnormal and positively skewed because excellent weather condition must prevail throughout the entire growing season if high yields are to be obtained while bad weather during any critical period can significantly reduce yields. However, positive skewness was found only for cotton and no significant skewness or negative skewness was found for corn and oats. Lognormality test results suggest that lognormal cannot be used for representing yield distributions in most cases. Day employed the Pearson system of density functions[2] to approximate unknown yield distributions. Despite its flexibility in representing various shapes of yield distributions, this system is restrictive for stochastic processes that have strongly skewed PDF's or PDF's with multiple modes.

Gallagher (1987) found negative skewness for U.S. average soybean yields and he reasoned, "Yield cannot exceed the biological potential of the plant, yet it can approach zero under blight, early frost or extreme heat". He also found that "soybean yield variability has changed systematically over the past five decades" and "the variance of U.S. soybean yields has been increasing". He modeled skewness and changing variance

[2] It was developed by Karl Pearson(1895). The Pearson system is determined by its first four moments.

of soybean yields using a gamma distribution function[3]. Norwood *et al.* (2004) pointed out that it is difficult to identify the maximum yield needed to implement the GAMMA model when conducting forecasts.

Nelson and Preckel (1989) used a conditional beta distribution as a priori assumption to model the distribution of farm-level corn yields in five Iowa counties and confirmed negative skewness. They estimated the maximum attainable yield by maximum likelihood and modeled deviations of yield from its maximum value as a conditional beta distribution. The beta distribution has the flexibility of fitting skewness in either direction and may have the bell-shape suggested by Day. Their analysis did not consider correlation of yields between farms in the same county when pooling farm-level data to estimate corn yield response to fertilizer applications.

Moss and Shonkwiler (1993) estimated a stochastic trend model with possible nonnormal disturbances for U.S. corn yields. They found negative skewness using an inverse hyperbolic sine transformation (IHST) of residuals. IHST simultaneously shrinks large residuals toward zero and parameterizes residual skewness and kurtosis within the stochastic trend framework. Homoskedasticity has to be imposed because the Kalman filter loses tractability under heteroskedasticity.

Goodwin and Ker (1998) used a flexible nonparametric kernel method to model county-level crop yield distributions in their study of rating group-risk crop insurance contracts.

[3] The density function of a gamma distribution can be expressed by two parameters.

Nonparametric kernel models make no explicit assumption about the functional form of the yield probability density function. They commented that, because kernel density estimation techniques require the choice of a kernel function and bandwidth, and the rate of convergence to the true density is relatively slow, a parametric specification would be more desirable if one knows the true functional form of the density.

Just and Weninger (1999) challenged the predominant view that crop yield distributions are nonnormal by arguing that rejection of normality may be the result of methodological and data limitations and the normal distribution should remain a reasonable candidate for modeling yield densities. They found that normality is difficult to reject when flexible polynomial trends are used for mean yield and yield variance.

Ramirez *et al.* (2003) addressed the procedural issues raised by Just and Weninger by using improved model specifications, estimation and testing procedures. Their findings reaffirmed nonnormality and left skewness of Corn Belt corn and soybean yields. They emphasized that because the type-two errors in the normality tests are unknown, nonrejection does not prove yield normality.

Ker and Coble (2003) proposed a semiparametric method for modeling yield distributions. Normal and beta distributions were used first to estimate corn yield densities and nonparametric kernel estimator was employed to correct the estimates. Their simulation results indicated that the semiparametric estimator with a normal distribution is more efficient than the competing parametric models (Normal and Beta) and the standard

nonparametric kernel estimator.

Sherrick *et al.* (2004) compared the beta, Weibull, logistic, normal and lognormal distributions in estimating farm corn and soybean yield densities and found that the Weibull and beta distribution ranked highest based on the goodness-of-fit measures. They also calculated the expected payouts to APH insurance with results suggesting significant differences across distributions in the expected value of APH insurance due solely to distributional assumptions.

Using the Ramirez data, Norwood *et al.* (2004) applied a new method of model evaluation based on the out-of-sample log-likelihood function values to six popular yield models and found that a model developed by Goodwin and Ker outperforms the competing models in forecasting out-of-sample county average yields.

CHAPTER 3

CANDIDATE YIELD MODELS

Various approaches to representing yield distributions can be segmented into two primary groups: parametric and nonparametric, depending on whether they appeal to a known parametric distribution or, alternatively, whether they use nonparametric techniques. Most yield distribution models are of a parametric nature. Under this approach, a specific parametric distribution is selected a priori and parameters of the distribution are estimated using observed yield data. The entire yield distribution is fully represented by the parameter estimates. Past parametric approaches to estimating yield distributions used the normal, lognormal, the Pearson system, beta, gamma, logistic, Weibull, inverse hyperbolic sine transformation (IHST) of normality, and a mixture of normal distributions. Some yield models used nonparametric/semiparametric approaches based on kernel estimating technique. Nonparametric kernel methods do not make any explicit assumption about the functional form of the yield probability density function (PDF) and thus "distribution free".

In this chapter, six candidate distributions are discussed including the normal, beta, Weibull, IHST, a mixture of normals, and kernel methods. Recently, the normal distribution has been frequently examined in agricultural risk management and crop insurance literature owing to Just and Weninger's strong defense. Beta, Weibull, IHST, and kernel methods all have their support in the literature, but no consensus has been reached regarding the most appropriate characterization of crop yield distributions. Bi-

polarity of crop yields was proposed recently by Goodwin and Ker, but few studies have examined a mixture of two normal distributions which could be a potential candidate for modeling crop yields.

3.1 Normal (Gaussian) Distribution

Just and Weninger argued that since normality can't be easily rejected under correct model specification and normality tests, normal distribution should remain a reasonable candidate for yield models.

The density function of a normal distribution with mean μ and standard deviation σ (where $-\infty < \mu < \infty$, $\sigma > 0$) is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty$$

The entire distribution depends on two parameters, μ and σ. The normal distribution is symmetric, bell shaped, and unbounded in the real number line R. Generally the likelihood of predicting a crop yield being less than zero is negligible.

3.2 Beta Distribution

Most studies used the beta distribution as a priori distribution when sufficient evidence of skewness and/or kurtosis was found in their yield data.

The density function of the beta distribution is

$$f(x) = \frac{(x-a)^{p-1}(b-x)^{q-1}}{B(p,q)(b-a)^{p+q-1}},\ a \le x \le b,\ \ p,q > 0$$

where *a* and *b* are the lower and upper bounds of the distribution respectively, p and q are the shape parameters, $B(p,q)$ is the beta function. The beta function has the fourmula:

$$B(p,q) = \int_0^1 t^{p-1}(1-t)^{q-1}\,dt$$

Beta distribution allows a wide range of skewness and kurtosis, and can be symmetric as well. The upper and lower bounds of the distribution are either specified or estimated in yield modeling applications.

3.3 Weibull Distribution

Weibull distribution has been proposed recently for modeling yields with very few applications in the agricultural economics literature.

The density function of the 2-paramer Weibull distribution has the form:

$$f(x) = \frac{\gamma}{\alpha}\left(\frac{x}{\alpha}\right)^{\gamma-1}\exp\left(-(x/\alpha)^{\gamma}\right),\ \ x \ge 0,\ \ \gamma,\alpha > 0$$

where γ is the shape parameter and α is the scale parameter.

Weibull distribution has many appealing properties, such as being bounded by zero and allowing for a wide range of skewness and kurtosis.

3.4 IHST Model

Moss and Shonkwiler applied inverse hyperbolic sine transformation technique in modeling U.S. corn yields. Their approach allows flexible specifications of yields and simultaneous estimation of central tendency and nonnormality. A standard stochastic trend model (Kalman filter) was used to estimate the central tendency (or the mean) of the distribution, and nonnormality within the stochastic trend was estimated by IHST that corrects both skewness and kurtosis. A limitation of this model is that Kalman filter loses tractability under heteroskedasticity.

Wang *et al.* (1998) proposed another method which is a variation of the model developed by Moss and Shonkwiler:

$$\ln y_t = \beta_0 + \beta_1 t + \beta_2 \ln y_{t-1} + u_t,$$

$$u_t = \frac{\sinh(\theta(v_t + \delta))}{\theta} \qquad (1)$$

$$v_t \sim N(0,\varsigma^2).$$

where lny_t is the logarithm of yield at time t; u_t is the non-normal error; v_t is an independently and identically distributed (i.i.d) normal disturbance with mean zero and variance ς^2; *sinh( )* is the hyperbolic sine transformation; δ and θ measure skewness and kurtosis respectively. The lagged term lny_{t-1} captures autocorrelation so that u_t is i.i.d. If $\delta>0$ ($\delta<0$), the distribution is skewed to the right (left) and if $\delta=0$ the distribution is symmetric. If $\theta \neq 0$, the distribution is kurtotic.

The first equation in (1) is a time-series econometric model using observations on yield itself. It describes that the mean of the current yield might have relation with time and last period's yield. The observed or realized yield y_t is a sum of the mean yield and the stochastic nonnormal shock, u_t. The second equation transforms the nonnormal shock u_t into normal shock v_t by the modified IHST.

Because the model is nonlinear in parameters, maximum log likelihood is used to estimate the parameters. The maximum log likelihood function is given in equation (2):

$$\underset{\beta_0,\beta_1,\beta_2,\varsigma^2,\theta,\delta}{MAX} \; L = -\frac{1}{2}\sum_{t=1}^{T}[\ln \varsigma^2 + \frac{v_t^2}{\varsigma^2} + \ln(\theta^2 u_t^2 + 1)], \tag{2}$$

$$v_t = \frac{\sinh^{-1}(\theta u_t)}{\theta} - \delta$$

$$= \frac{1}{\theta}\ln(\theta u_t + \sqrt{(\theta u_t)^2 + 1}) - \delta,$$

$$u_t = \ln y_t - \beta_0 - \beta_1 t - \beta_2 \ln y_{t-1}$$

Six parameters β_0, β_1, β_2, ς, θ and δ are chosen to maximize the log likelihood function.

3.5 A Mixture of Normals

Few empirical applications used a mixture of normal distributions to model yields. Ker and Goodwin (2000) argued that it is possible for the unknown yield distribution to be bi-modal and negatively skewed due to the effects of catastrophic events such as drought, flood and freeze. Observed yields can be seen as drawn from one of two distinct sub-populations: a catastrophic sub-population and a non-catastrophic sub-population. That is, if a catastrophic event occurs in a particular year, yields are drawn from the

catastrophic sub-population; if no such an event occurs in a year, yields are drawn from the non-catastrophic population. The distribution from catastrophic years (secondary distribution) lies on the lower tail of the distribution from non-catastrophic years (primary distribution) and has considerably less mass, which leads to negative skewness of yields. The secondary distribution lies to the left of the primary distribution because yields tend to be much lower in catastrophic years. It would also be expected that the secondary distribution has less mass since catastrophic events occur with much less frequency than their complement. Therefore, yield distribution could be negatively skewed and bi-modal if the mass of catastrophic distribution is non-negligible and the catastrophic distribution is relatively peaked.

Above structure proposed by Ker and Goodwin seems reasonable in terms of assuming the existence of two distinct sub-populations of yields. They considered yield distribution as being formed by primary distribution from non-catastrophic years and secondary distribution from catastrophic years, one on the right and the other on the left. This might cause confusion because given the two sub-populations the entire yield distribution is not a simple connection of the two distributions at some point. Instead, it is a mixture of the two distributions. In other words, density of any particular yield level is an expected density or, a mixture of the densities from two distributions.

The probability density function of a mixture of normal distributions has the following form:

$$f(y; \mathbf{p}, \boldsymbol{\mu}, \boldsymbol{\sigma}) = \sum_{i=1}^{k} p_i g(y; \mu_i, \sigma_i)$$

$$g(y; \mu_i, \sigma_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{(y - \mu_i)^2}{2\sigma_i^2} \right]$$

$$\sum_{i=1}^{k} p_i = 1$$

where y is the variable of interest; $g(y; \mu_i, \sigma_i)$ is the normal density function with mean μ_i and standard deviation σ_i; $p_1, \ldots, p_k$ are the mixing probabilities. The parameters $\mu_{1,\ldots,}\mu_k$, $\sigma_{1,\ldots,}\sigma_k$ and $p_{1,\ldots,}p_k$, are estimated by maximum likelihood along with standard errors obtained from the observed information matrix, i.e., the inverse of the Hessian matrix.

3.6 Kernel Density Estimator

Parametric approaches assume that yield distribution follows a known functional form, but specific functional form assumptions about crop yield distributions may not be well justified. Contrary to parametric estimators, nonparametric estimators do not assume any particular functional form for yield PDF.

Nonparametric kernel density estimators approximate the PDF from observations on the variable of interest and therefore are fully flexible in capturing local idiosyncrasies in yield distributions that may not be properly reflected in parametric specifications. In addition, kernel density estimators are mathematically tractable and easy to implement.

The nonparametric kernel density estimator is the vertical sum of densities placed over each observation. Unlike histograms, the individual kernel intervals, or windows, are

allowed to overlap. Rather than merely counting the number of observations in a window as in a histogram, a weight between 0 and 1 is assigned to each value based on its distance from the center of the interval and it is the weighted values that are summed. The kernel is a function determining these weights. A kernel density estimator has the form:

$$f(x) = \frac{1}{n \times h} \sum_{i=1}^{n} K\left[\frac{x - X_i}{h} \right]$$

where K is a symmetric probability density function satisfying the condition:

$\int_{-\infty}^{\infty} K[z]dz = 1$, X_i is the observation, n is the number of observations, and h is the window-width or the smoothing parameter. Kernel functions can be Gaussian, Epanechnikow, triangular, biweight, cosine, or rectangular density.

Silverman(1986) evaluated the efficiency of many potential kernels in terms of mean integrated squared error and concluded that, while there is little difference between the potential kernels, the Epanechnikov kernel is the most efficient in minimizing the mean integrated squared error. The Epanechnikov kernel that has the form:

$$K(z) = \begin{cases} \frac{3}{4\sqrt{5}}(1 - \frac{1}{5}z^2), if\,|z| \le \sqrt{5} \\ 0, otherwise. \end{cases}$$

The choice of window-width or smoothing parameter, h, controls the amount by which the data are smoothed in a kernel density estimate. The smoothing parameter is similar to the inverse of the number of bins in a histogram; smaller widths mean more detail.

The choice of the value of h is essentially a compromise between smoothing enough to remove insignificant bumps and not smoothing too much to smear out real peaks.

CHAPTER 4

STATIONARY YIELD DATA

Time-series crop yield data may be nonstationary, i.e., not generated by the same data generating process (DGP) due to technologic, economic, and behavioral changes. The trend component, in any (either deterministic or stochastic), must be controlled for before assessing the yield distribution at a point in time. In this chapter, trend estimation methods are discussed and two approaches for deriving stationary yield data are proposed.

4.1 Detrending Yield Data

Many studies have used a deterministic trend (for example, a simple linear trend) to model the growth path of average yield assuming that mean yield increases at a deterministic rate due to technology developments. However, technology generating/adopting function is not necessarily deterministic, or in other words, crop yield may grow at a random rate over time. If yield series has a stochastic trend, regressing yield against time functions would result in spurious parameter estimates. Therefore, stochastic trend should be tested before identifying any potential deterministic trend in the yield series. Augmented Dickey Fuller (ADF) and Phillips-Perron (PP) tests are the most common approaches for testing for the existence of a stochastic trend. The ADF test is used in this paper. If the unit root hypothesis can't be rejected, we should difference the yield series.

After controlling for the stochastic trend component (or passing the unit root test), deterministic trend should be examined. Just and Weninger suggested polynomial time functions for estimating deterministic trends. Some studies regressed yields against a polynomial time function and tested down towards the linear trend based on F-tests. Specifically, $y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + ... + \beta_n t^n + \varepsilon_t$ is assumed by arbitrarily choosing *n* and tested down based on F-tests.

This paper proposes a model that nests both trend and autocorrelation possibilities within a time-series model specification:

$$y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 y_{t-1} + u_t$$

where y_t is the yield (or differenced yield if there is a unit root) at time *t,* y_{t-1} is the one-period lagged yield, and u_t is the disturbance. Quadratic trend is assumed because empirical studies in this research area found evidence for at most quadratic trend and generally linear trend (Sherrick *et al.*). Yield data are fitted to above model and t-tests are used to determine the appropriate time-series model.

Next, adequacy of the estimated time-series model is investigated using diagnostic tests including tests for the presence of heteroskedasticity and serial correlation. Specifically, the presence of heteroskedasticity is assessed using White test and serial correlation using general regressors (Wooldridge).

After obtaining evidence of adequacy to the time-series model, yield data are detrended to a base year before assessing yield distributions. Two methods can be used to detrend

yield data based on different assumptions. If constant Coefficient of Variation (CV) is assumed, yield data are detrended to the base year according to:

$$y_t' = \hat{y}_{baseyear} \times \frac{y_t}{\hat{y}_t}$$

where $\hat{y}_{baseyear}$ is the predicted yield for the base year from the time-series model, $\frac{y_t}{\hat{y}_t}$ is the ratio of observed yield to predicted yield for period *t*, and y_t' is the detrended yield. If no assumption about the CV is made, yields are detrended using $y_t' = \hat{y}_{baseyear} + (y_t - \hat{y}_t)$. Through these methods, all yields are converted to the base year equivalents which are used to estimate candidate yield distributions.

Some studies employed IHST models to simultaneously capture yield trend and estimate the distribution of the residuals derived from the yield trend estimation. This paper applies IHST to a case study as one of the candidate yield models. ARIMA was used by Goodwin and Ker to model yield trends. Generally large samples are required for ARIMA model estimation. This paper does not apply this method to the case study due to the limited sample size.

CHAPTER 5

FARM-LEVEL AND COUNTY-LEVEL YIELD DISTRIBUTIONS

Yield volatility at the farm level is crucial for studying various crop insurance programs and production under risk and uncertainty. Unfortunately, historical farm-level yield data are available for relatively short time horizons (at most 20 years and generally much less), which poses a major obstacle in identifying the shape of farm-level yield distributions. Most studies resorted to county-level yield data and conjectured about farm-level yield distributions using county-level yield distributions. This method may lead to serious mischaracterization of farm-level yield risk due to the averaging effect. We investigate the relationship between farm- and county-level yield PDF's using both statistical theory and Monte-Carlo simulation study in this chapter.

5.1 Averaging Effect

Very few farm yield modeling applications estimated county-level yield PDF and converted it into the farm-level PDF using a mean preserving spread technique. These studies conjectured that yield distribution at the county level might be similar to that at the farm level as geographical conditions in a given county are similar across farms while the variance at the county-level is smaller than at the farm-level due to the averaging effect. Most research on crop yield distributions either just estimated county-level yield PDF's or implicitly assumed that farm-level yield PDF's can be represented by county-level yield PDF's. Surprisingly, the relationship between farm- and county-level yield PDF's has never been formally investigated in the literature. This chapter seeks to fill

this research gap.

Data generating process (DGP) for crop yields at the farm level is different from that at the county level. County yield is the *average* of correlated farm yields, for which there is no recognizable PDF as indicated by statistical theory. Thus using county-level yield PDF to conjecture about farm-level yield PDF is dangerous. It might lead to serious mischaracterization of farm-level yield distributions.

Farm-level and county-level yield PDF's do not share the same distributional form except for an extreme case when farm yields in a given county are *independent* and normally distributed. That is, if farm yields $y_i \sim Normal(\mu_i, \sigma_i)$, i=1,…, n, and y_i are independent , then the county yield, $\frac{1}{n}\sum_{i=1}^{n} y_i$, is also normally distributed with mean $\frac{1}{n}\sum_{i=1}^{n} \mu_i$ and standard deviation $\frac{1}{n}\sqrt{\sum_{i=1}^{n} \sigma_i^{\,2}}$, where n is the number of farms in the county, y_i is the yield for farm i, μ_i and σ_i are the mean and standard deviation for farm i respectively. Another important statistical theory, central limit theorem, indicates that county yields are approximately normally distributed if there is sufficient number of farms in a given county and farm yields are *independent* (they can have any PDF). Under central limit theorem county and farm yield distributions may have very different shapes, so inferring about farm-level distribution based on county-level distribution can be misleading if conditions for the central limit theorem are met. Nevertheless, these theorems regarding independent random variables are generally not applicable to the real

world because farm yields in a given county are often correlated due to similar weather conditions and soil types.

As long as farm yields in a county are correlated, there is no recognizable PDF for the county yields. However, we suspect that the structure of yield PDF's at the farm and county levels may be similar in some cases when non-systemic randomness is relatively weak compared to systemic randomness. We use Monte-Carlo simulation method to investigate the relationship between farm- and county-level yield distributions when farm yields are correlated.

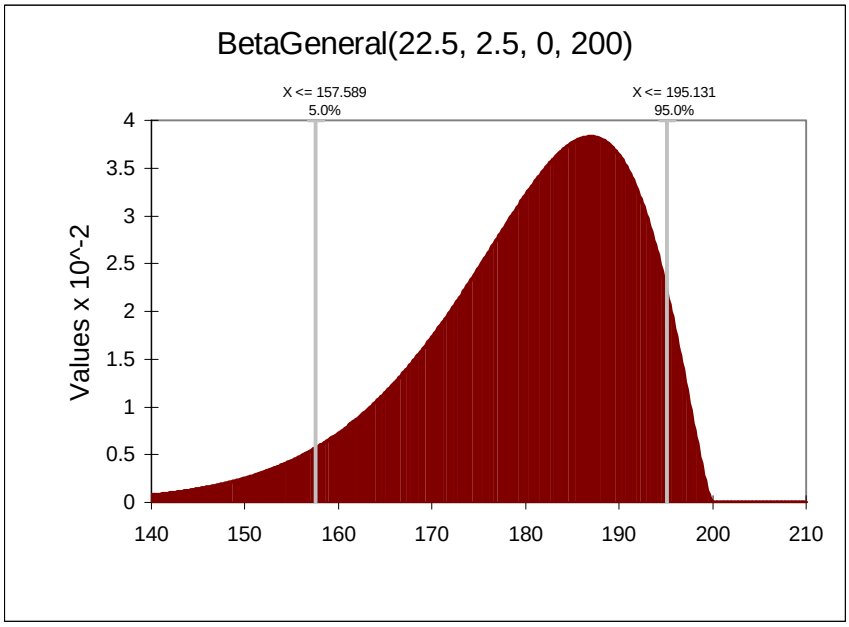
5.2 Monte-Carlo Simulation Analysis

A simple Monte-Carlo simulation is conducted using @Risk software as follows:

(1) Set the number of farms in a county: n=5

(2) Specify the functional form of farm-level yield PDF's: BetaGeneral(22.5, 2.5, 0, 200) is chosen for all the farms for simplicity, where 22.5 and 2.5 are the shaper parameters, 0 is the lower bound and 200 is the upper bound. The PDF graph of this distribution is given in Figure 1.

Figure 1: Beta Distribution



(3) Specify the correlation matrix between farm yields:

	Farm A	Farm B	Farm C	Farm D	Farm E
Farm A	1	0.8	0.8	0.8	0.8
Farm B	0.8	1	0.7	0.7	0.6
Farm C	0.8	0.7	1	0.6	0.7
Farm D	0.8	0.7	0.6	1	0.7
Farm E	0.8	0.6	0.7	0.7	1

(4) Set county yield as the average of farm yields and run the simulation

(5) Obtain ranking of candidate fits for the simulated county yield data based on Chi-square, Anderson-Darling and Kolmogorov-Smirnov statistics.

The beta distribution is ranked highest among common parametric distributions. Test statistics for the first three best fits are presented below in Table 1.

Table 1: Distribution Fits for Simulated County Yields

		Beta	Logistic	Normal
Chi-Sq	Test Value	34.47	754.38	793.98
	P Value	0.99	0	0
	Rank	1	2	3
A-D	Test Value	0.25	46.02	64.17
	P Value	not available	< 0.005	< 0.005
	Rank	1	2	3
K-S	Test Value	0.01	0.06	0.07
	P Value	not available	< 0.01	< 0.01
	Rank	1	2	3

Fitted distributions are plotted in Figures 2 – 4.

Figure 2: Fitted Beta Distribution

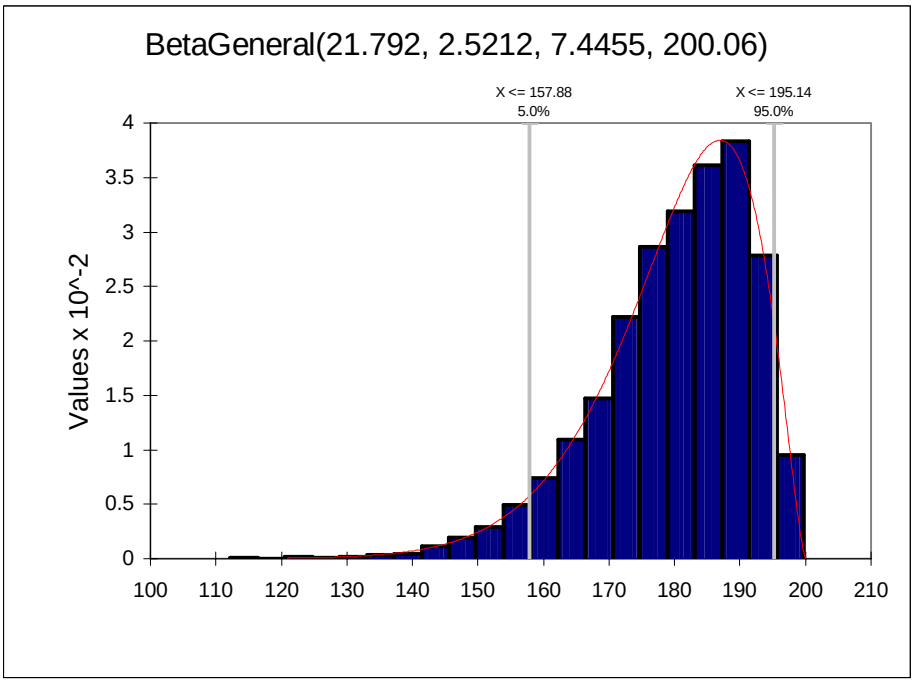


Figure 3: Fitted Logistic Distribution

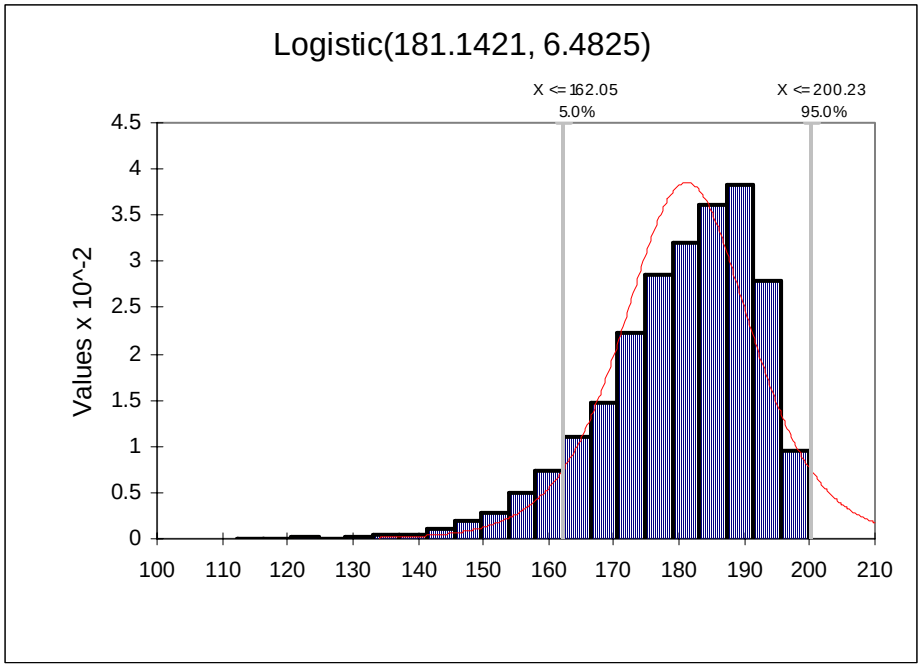
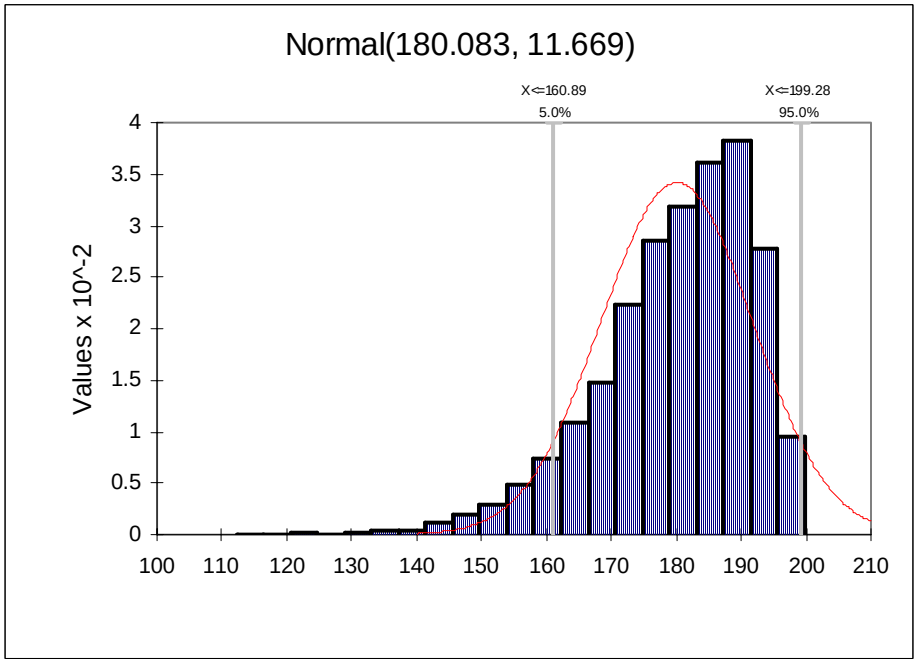


Figure 4: Fitted Normal Distribution



Above simulation results indicate that a Beta distribution fits the simulated county yield data very well and much better than other functional forms (based on P-values) although theoretically county yields can not be strictly characterized as a Beta distribution. Hence we might be able to conclude that county-level and farm-level yield distributions are similar in shape under some circumstances such as high yield correlations and similar structure of yield PDF's across farms.

Ker and Coble stated in their footnote, "Our results are with respect to county crop-yield distributions and using them to conjecture about farm-level distributions is dangerous. Having given sufficient warning, we feel there are cases where these conjectures are more reasonable than others. As spatial area increases, spatial averaging reduces the effect of non-systemic or local randomness while the systemic or area randomness remains. However, depending on the degree of systemic randomness versus non-systemic randomness for the crop-farm combination, the structure of yield densities at the county and farm level may be quite similar. In instances where the non-systemic randomness is relatively weak, for example, where GRP is in demand, our results could shed light on the structure of farm-level yield distributions." Their statement is supported by the formal simulation analysis of this chapter.

This section examines the relationship between county- and farm-level yield distributions under one Monte-Carlo simulation setting. Future research in this area would investigate broader scenarios including alternative number of farms, multiple candidate functional forms and correlation structures.

CHAPTER 6

A CASE STUDY OF CORN YIELDS IN BRANCH COUNTY, MICHIGAN

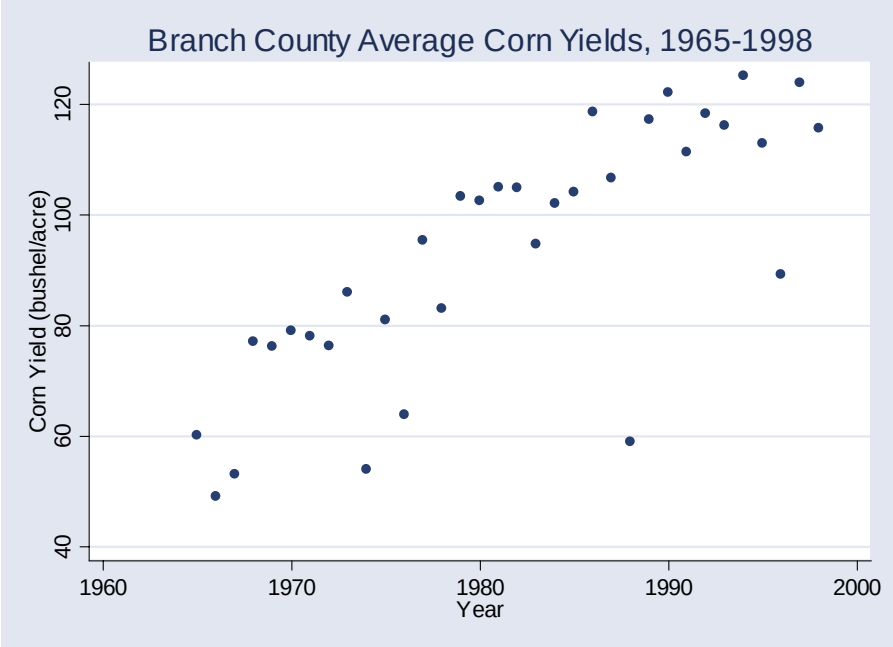
In this chapter, we conduct a case study of estimating corn yield PDF for Branch County, Michigan. Section 1 detrends corn yield data; Section 2 fits Normal, Beta, and Weibull distributions to the detrended yield data and compares these distributions using goodness-of-fit results; Section 3, 4, and 5 estimate IHST, a mixture of two normals and kernel density respectively.

6.1 Detrending Yield Data

Annual corn yield data for Branch County, Michigan were collected from the National Agricultural Statistics Service (NASS) of the United States Department of Agriculture.

Historical corn yields are plotted in Figure 5.

Figure 5: Branch County Average Corn Yields (1965-1998)



Summary statistics are given in Table 2.

Table 2: Summary Statistics for Branch County Corn Yields (1965-1998)

Variable	Mean	Standard Deviation	Minimum	Maximum
Corn yield	93.03	22.74	49.00	125.11

Number of Observations=34

Figure 5 suggests an upward trend in the average corn yields. Dickey-Fuller unit root test was conducted first to examine if there is any stochastic trend component in the yield series. The following model was estimated for the augmented Dickey-Fuller test:

$$\Delta cornyield_t = \beta_0 + \beta_1 \cdot t + \beta_2 \cdot cornyield_{t-1} + \sum_{i=1}^{2} \gamma_i \cdot \Delta cornyield_{t-i} + u_t,$$

where $\Delta cornyield_t$ is the first difference of $cornyield_t$, t is the time trend, $cornyield_{t-1}$ is the lagged $cornyield_t$, and $\Delta cornyield_{t-i}$ is the ith lag of $\Delta cornyield_t$.

The two lags of $\Delta cornyield_t$ were both individually and jointly insignificant at the 5% level, so they were dropped from the model and the Dickey-Fuller test was conducted again for

$$\Delta cornyield_t = \beta_0 + \beta_1 \cdot t + \beta_2 \cdot cornyield_{t-1} + u_t$$

The estimate of β_2, coefficient on $cornyield_{t-1}$, is –1.01 and the test statistic is –5.498, so we reject the null hypothesis unit root at the 5% level. OLS results of the two regressions are presented in Table 3.

Table 3: OLS Results

Dependent Variable: $\Delta cornyield_t$

t	2.001 (0.714) [0.525]	1.80 (0.43)
$cornyield_{t-1}$	-1.22 (0.347) [0.274]	-1.01 (0.18)
$\Delta cornyield_{t-1}$	0.147 (0.276) [0.227]	___
$\Delta cornyield_{t-2}$	0.097 (0.205) [0.101]	___
constant	79.2 (20.84) [16.2]	62.26 (12.11)
Observations	31	33
R-squared	0.5505	0.5023
Adj. R-squared	0.4813	0.4691

Note: The quantities in parentheses below the estimates are the standard errors, and in brackets are the robust standard errors.

Now we can safely use the yield data to estimate the following time-series model:

$$cornyield_t = \beta_0 + \beta_1 \cdot t + \beta_2 t^2 + \beta_3 \cdot cornyield_{t-1} + u_t$$

The estimated equation is:

$$\widehat{cornyield}_t = 54.92 + 3.74t - 0.05t^2 - 0.08cornyield_{t-1}$$
$$(12.69)\ (1.29)\ (0.03)\ \ (0.18)$$
n=33 R-squared=0.6331 Adj. R-squared=0.5951

Coefficient estimate on $cornyield_{t-1}$ has a t statistic of –0.438, which is statistically insignificant at the 5% level, so $cornyield_{t-1}$ was dropped from the model. The equation is estimated to be:

$$\widehat{cornyield}_t = 52.47 + 3.296t - 0.04t^2$$
$$(7.55)\ \ (0.994)\ (0.03)$$
n=34 R-squared=0.6534 Adj. R-Squared=0.6310

For the above model, we conducted diagnostic tests, i.e., we tested for heteroskedasticity and serial correlation. The F statistic from the White test for the heteroskedasticity is 0.25 and the p-value is 0.78, so we can't reject the null hypothesis of homoskedasticity. The test for serial correlation with general regressors gives the t statistic of -0.44 and the p-value of 0.67, which implies no serial correlation in the residuals. The results from diagnostic tests indicate that the above time-series model is adequate.

The model was re-estimated using robust regression method[4] that weights observations automatically so that the influence of outliers can be reduced.

Table 4 shows the robust regression result of the model:

$$cornyield_t = \beta_0 + \beta_1 \cdot t + \beta_2 t^2 + u_t$$

Table 4: Robust Regression Results

	Coefficient	Std. error	t statistic	P-value
t	3.73	0.63	5.88	0.000
t^2	-0.052	0.018	-2.94	0.006
Constant	52.34	4.81	10.88	0.000

Number of obs.=34, F(2,31)=82.37, Prob>F=0.0000

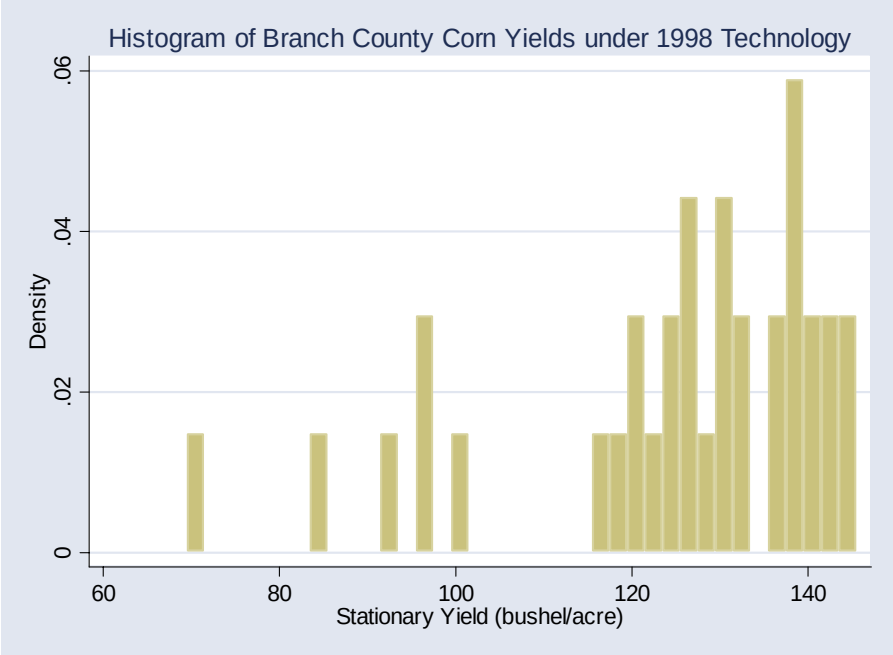
Next yield data were detrended to the base year of 1998 using

$$cornyield'_t = \hat{cornyield}_{1998} + (cornyield_t - \hat{cornyield}_t)$$

Histogram of the stationary yield data is displayed in Figure 6.

[4] See appendix.

Figure 6: Histogram of Branch County Corn Yields under 1998 Technology



6.2 Normal, Beta and Weibull Estimates and Goodness-of-fit Tests

Using the detrended yield data from section 6.1, this section estimates Normal, Beta and Weibull distributions by maximum likelihood estimation method and compares these distributions by goodness-of-fit results including Chi-square, Anderson-Darling and Kolmogorov-Smirnov statistics[5].

Figures 7 – 9 display the parameter estimates and PDF plots for the estimated distributions.

[5] @RISK software is used to estimate candidate distributions in this section.

Figure 7: Normal Estimates

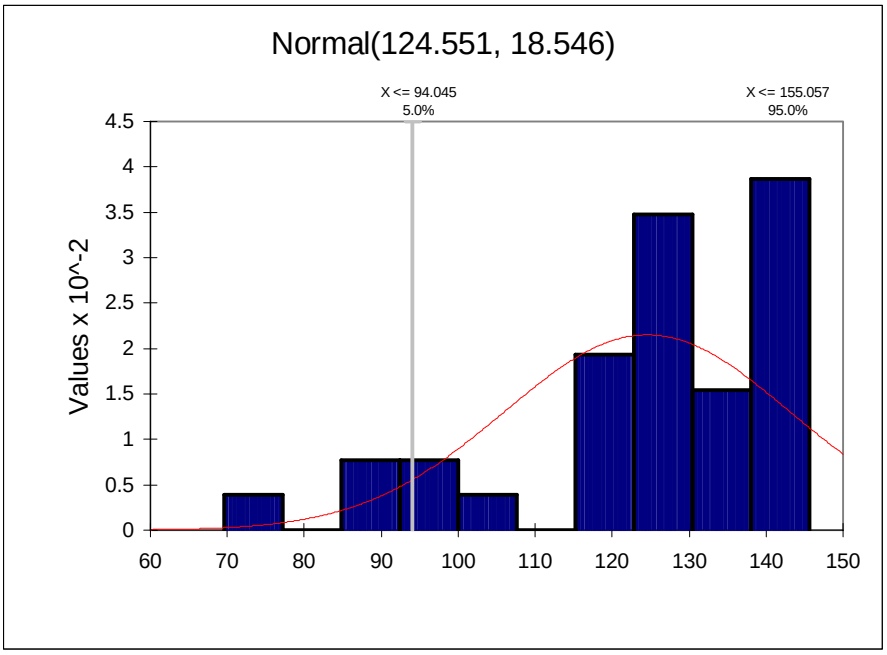


Figure 8: Beta Estimates

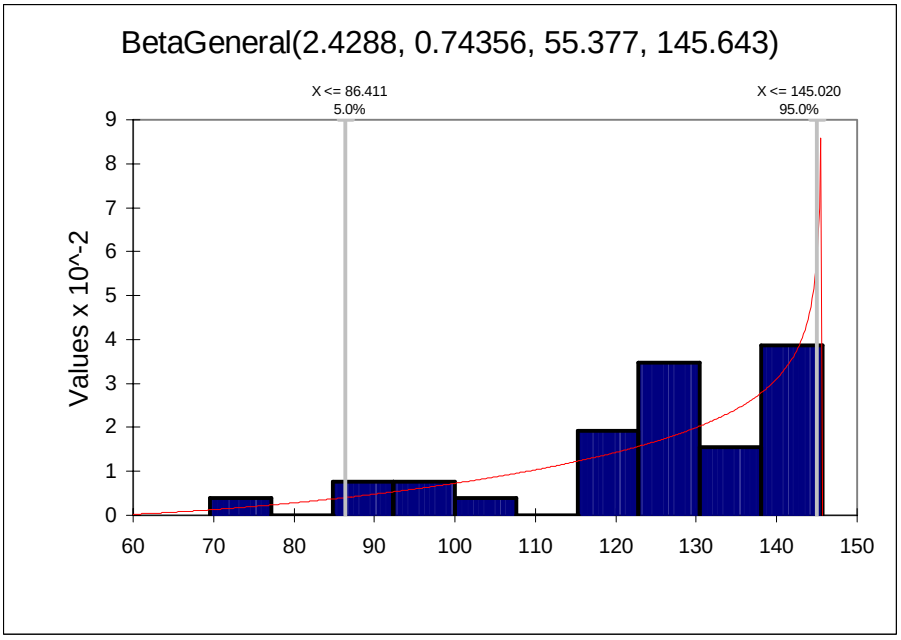
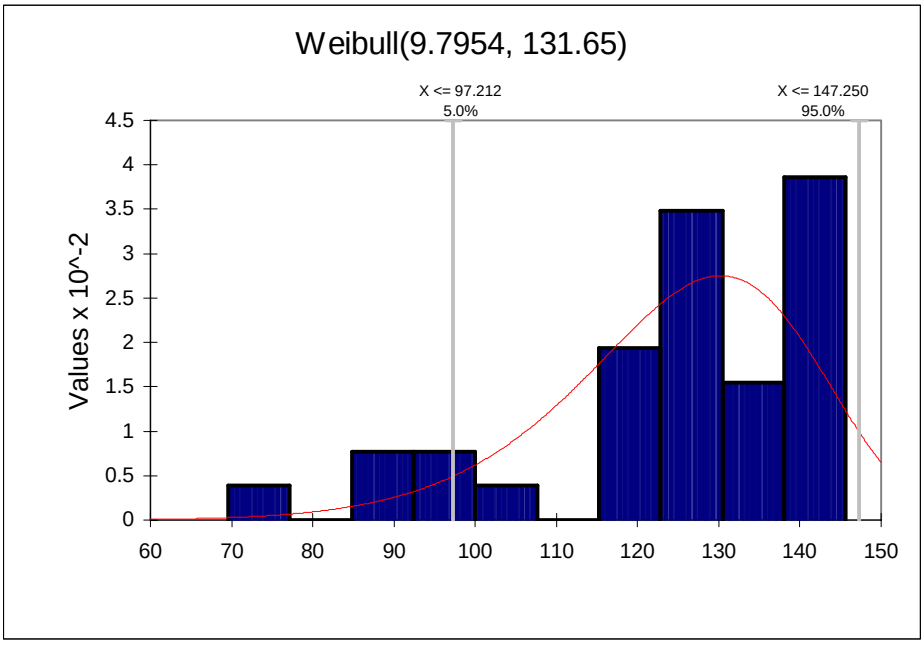


Figure 9: Weibull Estimates



Goodness-of-fit tests for Normal, Beta, and Weibull distributions are reported in Table 5.

Table 5: Goodness-of-Fit Tests of Candidate Distributions

		Normal	Beta	Weibull
Chi-Sq	Test Value	14.18	8.41	4.71
	P Value	0.0277	0.2095	0.5821
A-D	Test Value	1.65	Infinity	0.98
	P Value	< 0.005	N/A	0.01 <= p <= 0.025
K-S	Test Value	0.18	0.11	0.12
	P Value	< 0.01	N/A	> 0.1

Normality of yield distribution is strongly rejected by Chi-square, Anderson-Darling (A-D) and Kolmogorov-Smirnov (K-S) tests, which indicates that the corn yield distribution is very unlikely to be normal. Weibull distribution is rejected by A-D test, but not rejected by Chi-square and K-S tests at the 5% level; Beta distribution is not rejected by the Chi-square test, but the P-values of A-D and K-S tests for Beta distribution are not

available. The results from Chi-square tests suggest that the Weibull distribution fits the data best compared to the normal and beta distributions because it has the largest P-value.

6.3 IHST Estimates

This section estimates the IHST model which has the form:

$$\log(cornyield_t) = \beta_0 + \beta_1 \cdot t + \beta_2 t^2 + u_t,$$

$$u_t = \frac{\sinh(\theta(v_t + \delta))}{\theta}$$

$$v_t \sim N(0,\varsigma^2)$$

where u_t is the non-normal disturbance; *sinh( )* is the hyperbolic sine transformation (HST); v_t is an i.i.d normal disturbance with mean zero and variance ς^2; δ and θ are parameters measuring skewness and kurtosis respectively. Generally, when δ is positive u_t is skewed to the right, if it is negative u_t is skewed to the left, and if it is zero u_t is symmetric. When θ is zero, u_t is as kurtotic as the normal distribution (the limit of u_t is $v_t+\delta$ as θ approaches zero) and u_t becomes more and more kurtotic as the magnitude of θ increases either positively or negatively.

The first equation is the trend equation, which describes the central tendency of the current yield as being determined by time. The transformed stochastic variable, u_t, has a non-zero mean which is determined by the three parameters of the transformation. Although still deterministic, the central tendency of $cornyield_t$ is the sum of the first three terms and $E(u_t)$ in the equation. The realized yield is a combination of the deterministic central tendency and a stochastic non-normal shock. The second equation converts the

non-normal shocks to normal shocks by the modified HST. Because the model is nonlinear in parameters, maximum log likelihood estimation method is used for estimating the parameters[6]. The maximum log likelihood function is:

$$MAX\,L = -\frac{1}{2}\sum_{t=1}^{T}[\ln \varsigma^2 + \frac{v_t^2}{\varsigma^2} + \ln(\theta^2 u_t^2 + 1)]$$

$$v_t = \frac{\sinh^{-1}(\theta u_t)}{\theta} - \delta = \frac{1}{\theta}\ln(\theta u_t + \sqrt{(\theta u_t)^2 + 1}) - \delta$$

$$u_t = \log(cornyield_t) - \beta_0 - \beta_1 \cdot t - \beta_2 t^2$$

Estimation results are displayed in Table 6.

Table 6: IHST Model Estimates

Coefficient	Estimates and Standard Errors	t-statistic	P-value
β_0	4.0 (0.09)	42.85	0.00
β_1	0.044 (0.012)	3.55	0.001
β_2	-0.0007 (0.0003)	-1.93	0.063
θ	7.46 (3.66)	2.04	0.041
δ	0.014 (0.02)	0.718	0.472
ς	0.11 (0.02)	4.868	0.000

Kurtosis parameter θ is estimated to be positive and significant at 5% level. The estimate of the skewness parameter δ is positive and statistically insignificant. Other parameter estimates are significant at the 1% level except for β_2 estimate which is significant at 10%

[6] STATA software is used to estimate IHST in this section.

level. Insignificant skewness parameter estimate might be the consequence of applying MLE method to a small sample. It is know that although maximum likelihood estimators have many desirable large sample properties[7], they can be heavily biased for small samples. That is, using small samples for maximum likelihood estimation can result in convergence failures, improper solutions and low accuracy of parameter estimates and standard errors.

6.4 A Mixture of Two Normals Estimates

The probability density function of a mixture of two normal distributions has the form:

$$f(yield_i; \boldsymbol{p, \mu, \sigma}) = p_1 g(yield_i; \mu_1, \sigma_1) + p_2 g(yield_i; \mu_2, \sigma_2)$$

where $g(yield; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(yield - \mu)^2}{2\sigma^2} \right]$, and p_1, p_2 are the mixing probabilities.

Using the stationary yield data, the parameters p_1, p_2, μ_1, μ_2 and σ_1, σ_2 are estimated by maximum likelihood method. Standard errors are obtained from the observed information matrix, i.e., the inverse of the Hessian matrix[8].

Parameter estimates are reported in Table 7.

Table 7: A Mixture of Two Normals Estimates

Number of observations=34		Log Likelihood =-134.33
	Coefficient	Standard Error
μ_1	94.41	13.13

[7] The estimators are asymptotically consistent, unbiased and efficient, and the estimates are normally distributed if the sample is large enough

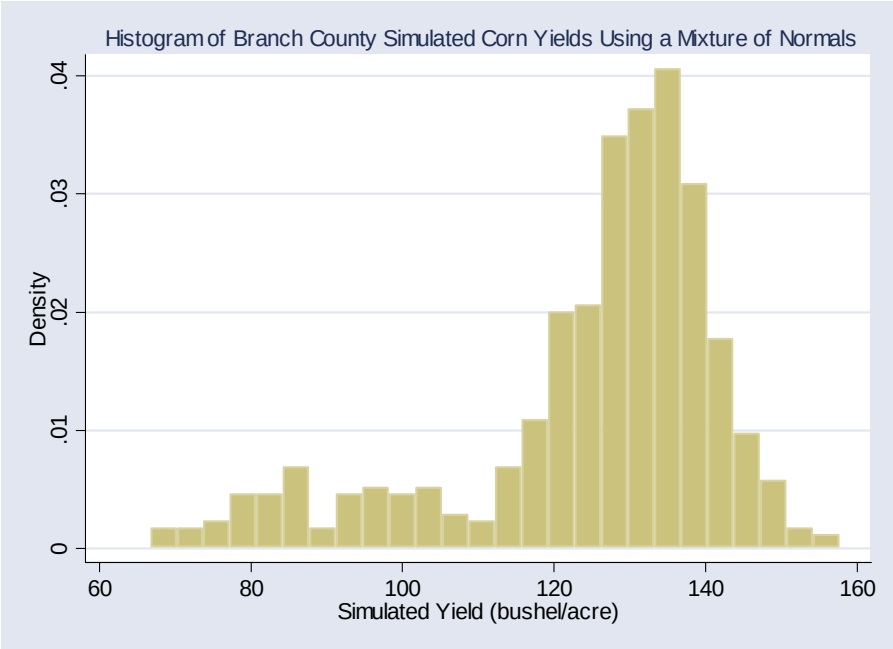
[8] STATA is used to estimate a mixture of two normals and simulate yields in this section.

μ_2	122.07	1.75
σ_1	17.84	6.67
σ_2	6.97	1.49
p_1	0.256	0.15

All parameter estimates are statistically significant at 5% level.

Histogram of simulated yields using these parameter estimates is shown in Figure 10.

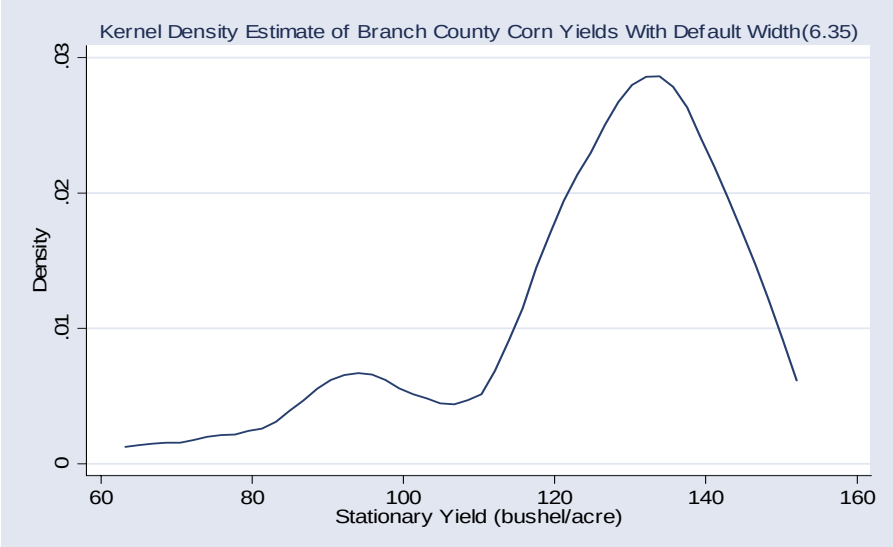
Figure 10: Histogram of Simulated Corn Yields Using a Mixture of Two Normal Estimates



6.5 Kernel Density Estimates

Using the stationary yield data, kernel density estimates of the yield distribution were obtained with STATA software. The graph of kernel density estimates with the default width in STATA is displayed in Figure 11.

Figure 11: Kernel Density Estimate of Branch County Corn Yields with Default Width



It is known that the default width is not necessarily the best. Multiple widths were specified and corresponding density graphs were inspected. Figures 12 and 13 display two of them.

Figure 12: Kernel Density Estimate of Branch County Corn Yields with 1.5×Default Width

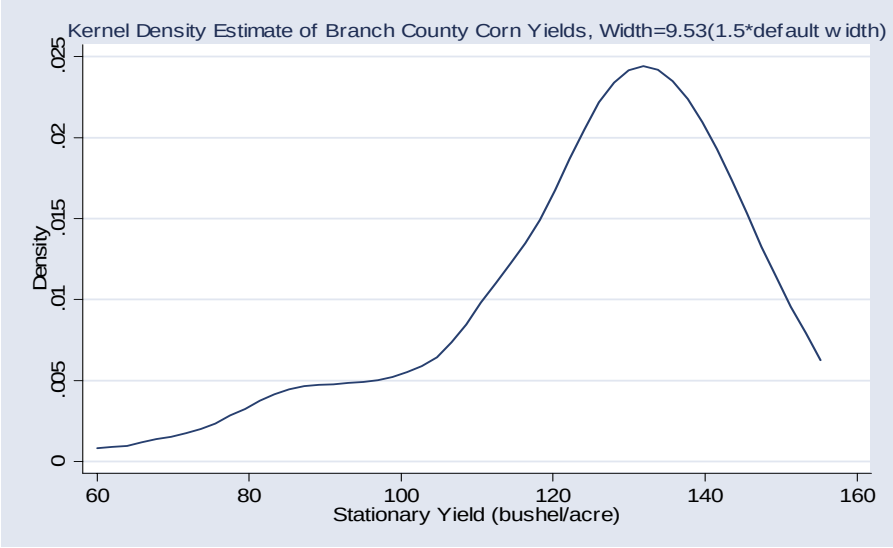
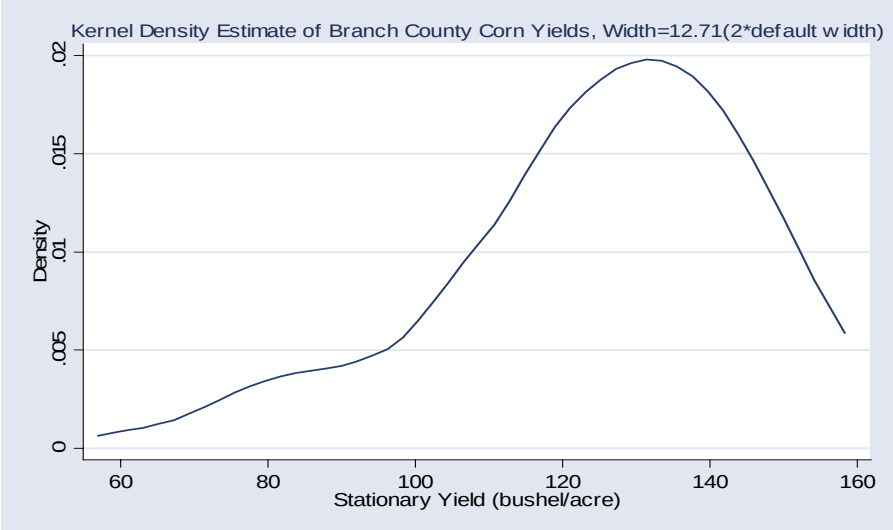


Figure 13: Kernel Density Estimate of Branch County Corn Yields with 2×Default Width



The window-width of 1.5×*default width* might be more appropriate compared to other widths that seem to either undersmooth or oversmooth the density graph.

The lower and upper bounds of yield distributions can be estimated from the kernel function. Epanechnikov kernel function is used in this section and its upper bound is $Xn+\sqrt{5}h$ and lower bound is $X_1-\sqrt{5}h$, where Xn and X_1 denote respectively the maximum and minimum observed yields, and h is the window width. For the Branch corn yields, $h=1.3\times r(width)=1.3\times 4.82=6.27$ and the end points are $62.79-\sqrt{5}\times 6.27=48.78$ and $138.48+\sqrt{5}\times 6.27=152.5$.

CHAPTER 7

SIMULATING FARM-LEVEL CORN YIELDS

This chapter proposes a new method for simulating farm-level yields using both county-level yield distribution estimates and information contained in farm-level yield data. The method is illustrated using a specific example.

7.1 Simulation Procedure

Results from Chapter 5 of this paper indicate that farm-level and county-level yield distributions can have similar structure, justifying the use of county yield PDF for assessing the shape of farm yield PDF when farm yield sample is not large enough. To derive farm yield PDF, some adjustments to the county yield PDF is warranted because of the potential differences in moments of farm- and county-level yield distributions such as mean and variance. This paper proposes that sample mean and variance of stationary farm yield data be used to adjust county-level PDF in estimating farm-level PDF. The same method is used to simulate farm-level crop yields in this section[9].

The simulation procedure is illustrated using kernel density estimates of Branch County corn yield distribution. First, the density estimates are converted to cumulative probabilities. Table 8 shows the data points, kernel density estimates, and cumulative probabilities which can be calculated in @RISK by the formula shown in Table 8.

[9] @RISK software is used for the simulation.

Table 8:@RISK Data Inputs

	A	B	C	
	Point	Density	Cumulative Probability	Formula for CDF
1	48.78	0	0	
2	56.52111	0.001259	0.004872	=1/2*(A2-A1)*B2
3	59.19473	0.00147	0.008519	=C2+1/2*(B2+B3)*(A3-A2)
4	61.86835	0.001567	0.012578	=C3+1/2*(B3+B4)*(A4-A3)
5	64.54196	0.002031	0.017387	
…	…	…	…	
35	144.7505	0.001259	0.993859	
36	152.5	0	0.998736	=C35+1/2*(A36-A35)*B35

Next, the mean and standard deviation (SD) of stationary farm yields for a farm in Branch County are assumed to be 76% of the county mean and 157% of the county SD[10]. Let Y denote the farm yield random variable and X denote the county yield random variable, then we have the equation $Y=0.76\mu_x+1.57(X-\mu_x)$, where μ_x is the expected value of X, which is set to be the sample mean of stationary county yields, 114.99 for Branch County. The above equation satisfies $E(Y)= 0.76\mu_x$ and $Var(Y)=1.57^2Var(X)$, where $E(Y)$ is the expected value of Y, $Var(Y)$ is the variance of Y, and $Var(X)$ is the variance of X.

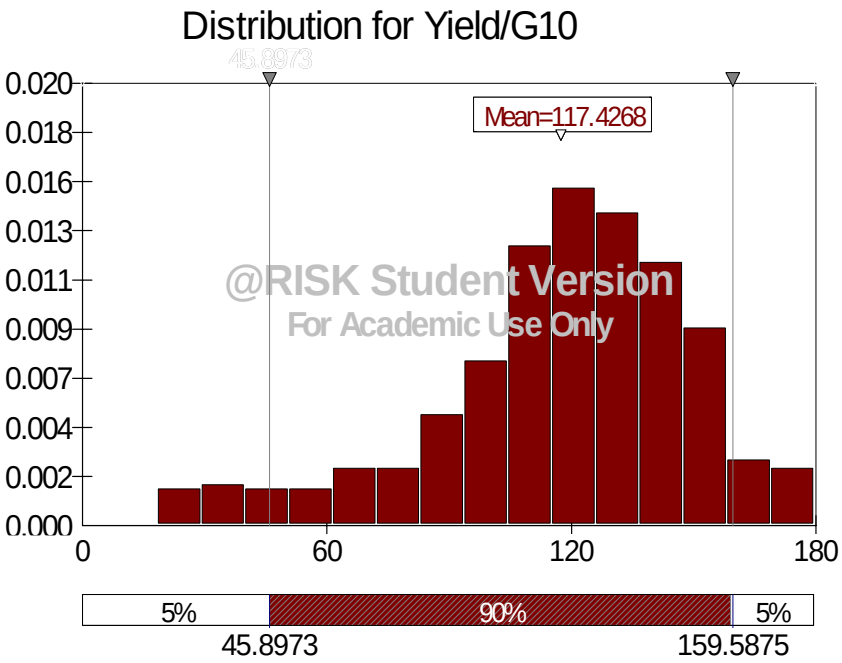
Using @RISK, The simulation input is specified as

0.76×114.99+1.57×(RiskCumul(A1,A36,A2:A35, C2:C35)-114.99).

The histogram of simulated farm corn yields is displayed in Figure 14.

[10] These values can be calculated using stationary farm and county yield data if farm data are available.

Figure 14: Histogram of Simulated Farm Corn Yields



CHAPTER 8

CONCLUSION

This paper expands the existing literature by suggesting methodological improvements in estimating and simulating farm-level yield distributions when farm-level yield data are insufficient. The results from Monte-Carlo simulation study of the relationship between farm-level and county-level yield distributions suggest that the yield density structure at the county level is similar to that at the farm level when farm yields are highly correlated and the yield density structure is similar across farms. However, enough caution should be exercised when inferring about farm-level yield distributions from county-level yield distributions because there are cases when the yield density structures at the farm level and county level can be quite different.

The results from the case study suggest that Branch County corn yields covered by the estimation period do not have a stochastic trend component; corn yields are adequately represented by a quadratic trend. There is no evidence of heteroskedasticity and serial correlation from the diagnostic tests.

The goodness-of-fit results indicate that normality of detrended yields is strongly rejected. Weibull distribution fits the data best compared to the normal and beta distributions based on the Chi-square statistics. Parameter estimates of a mixture of two normals are statistically significant and the model fits the data well, but bi-polarity of crop yield distributions needs to be further investigated in future study when more yield

data are available. The skewness parameter estimate of the IHST model is insignificant, which might be the consequence of applying MLE to a small sample size. The method for simulating farm-level yield distribution is illustrated using kernel density estimates. Future research would assess the economic importance of alternative yield distribution specifications on crop insurance policy rating and farm risk management decisions.

APPENDIX

ROBUST REGRESSION

Robust estimation is one that is insensitive to violations of any of the assumptions made about the way in which the data are generated. A lot of robust regressions were developed to deal with outliners included in the observations. STATA provides one version of robust regression using *rreg* command. It uses the iterated MAD[1] scale estimates, firstly performing an initial screening based on Cook's distance >1 to estimate gross outliers prior to calculating starting values and then performs Huber iterations followed by biweight iterations. Iterations stop when the maximum change in weights drops below the tolerance or the default value 0.01. Weights are derived from Huber weights function and Biweight function. Huber weights are used firstly until convergence and then biweights are used based on the result until convergence.

In Huber weighting, cases with small residual receive weights of 1 while those with larger residuals receive gradually smaller weights. Huber estimation obtains case weights by: $w_i = \begin{cases} 1, if \left|u_i\right| < c_h \\ c_h, otherwise \end{cases}$, where c_h is 1.345. So downweighting begins with cases whose absolute residual exceed (1.345/0.6745)MAD about or 2MAD. While, in biweights, all cases with non-zero residuals receive some downweighting, according to the smoothly decreasing biweight function:

[1] Let $e_1=y_i-x_ib$ represents the i_{th}-case residual. The i_{th} scaled residual $u_i=e_i/s$ calculated, where $s=MAD/0.6745$ is the residual scale estimate and $MAD=median(abs(e_i-median(e_i))$ is the median absolute deviation from the median residual.

$w_i = \begin{cases} (1-(u_i/c_b)^2)^2, if\,|u_i| < c_b \\ 0, otherwise \end{cases}$, where c_b=4.685/7*biweight tuning constant. The default value for biweight tuning costant is 7, which means in default, cases with absolute residuals of (4.685/0.6745)MAD or more are assigned to 0 wieghts and thus efficiently dropped.

BIBLIOGRAPHY

Day, R. H. "Probability Distribution of Field Crop Yields" *Journal of Farm Economics* 47(1965):713-41.

Gallagher, P. "U.S. Corn Yield Capacity and Probability: Estimation and Forecasting with Nonsymmetric Disburbances." *North Central Journal of Agricultural Economics* 8(1986): 109-22.

_____. "U.S. Soybean Yields: Estimation and Forecasting with Nonsymmetric Disturbances." *American Journal of Agricultural Economics* 69(1987): 796-803.

Goodwin, B. K., and Ker, A. P. "Nonparametric Estimation of Crop Yield Distributions: Implications for Rating Group-Risk Crop Insurance Contracts" *American Journal of Agricultural Economics* 80(1998): 139-53.

Hardle, W. *Applied Nonparametric Regression.* Econometric Society Monographs, 1990.

Just, R. E., and Weninger, Q. "Are Crop Yields Normally Distributed?" *American Journal of Agricultural Economics* 81(1999): 287-304.

Ker, A. P., and Goodwin, B.K. "Nonparametric Estimation of Crop Insurance Rates Revisited." *American Journal of Agricultural Economics* 83(2000): 463-78.

Ker, A.P. and Cobble, K. "Modeling Conditional Yield Densities" *American Journal of Agricultural Economics* 85(2003): 291-304.

Moss, C.B., and Shonkwiler, J.S. "Estimating Yield Distributions with a Stochastic Trend and Nonnormal Errors", *American Journal of Agricultural Economics* 75(1993): 1056-62.

Nelson, C.H. "The Influence of Distributional Assumption on the Calculation of Crop Insurance Premia." *North Central Journal of Agricultural Economics* 12(1990): 71-8.

Nelson, C. H., and Preckel, P. V. "The Conditional Beta Distribution As a Stochastic Production Function." *American Journal of Agricultural Economics* 71(1989): 370-78.

Norwood, B., Roberts, M.C., and Lusk, J. L. "Ranking Crop Yield Models Using Out-of-Sample Likelihood Functions" *American Journal of Agricultural Economics* 86(2004): 1032-1043.

Ramirez, O.A., Misra, S., and Field, J. "Crop Yield Distributions Revisited." *American Journal of Agricultural Economics* 85(2003): 108-20.

Silverman, B.W. *Density Estimation for Statistics and Data Analysis*. London: Chapman & Hall, 1986.
Wang, H. *Farmer Risk Management Behavior and Welfare under Alternative Portfolios of Risk Instruments.* Dissertation for the Ph.D. degree, Michigan State University, 1996.

Wang, H. H., Hanson, S. D., Myers, R. J., and Black, J. R. "The Effects of Crop Yield Insurance Designs on Farmer Participation and Welfare." *American Journal of Agricultural Economics* 80(1998): 806-20.

Wooldridge, J. M. Introductory Econometrics: A Modern Approach. South-Western College, 2000.