Geography of Crop Yield Skewness

Xiaodong Du
Dept. of Ag. & Applied Economics
University of Wisconsin-Madison
xdu23@wisc.edu

Cindy L. Yu
Dept. of Statistics
Iowa State University
cindyyu@iastate.edu

David A. Hennessy
Dept. of Economics
Iowa State University
hennessy@iastate.edu

Ruiqing Miao
Dept. of Economics
Iowa State University
miaorong@iastate.edu

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's 2012 AAEA Annual Meeting, Seattle, Washington, August 12-14, 2012

Geography of Crop Yield Skewness

Xiaodong Du
Assistant Professor
Dept. of Ag. & Applied Economics
University of Wisconsin-Madison
xdu23@wisc.edu

Cindy L. Yu
Associate Professor
Dept. of Statistics
Iowa State University
cindyyu@iastate.edu

David A. Hennessy
Professor
Dept. of Economics
Iowa State University
hennessy@iastate.edu

Ruiqing Miao
Post-doctoral Researcher
Dept. of Economics
Iowa State University
miaorong@iastate.edu

Abstract

This study seeks to provide a rigorous theoretical and empirical understanding of the effects of exogenous geographic and climate-related factors on the first three moments of crop yields. We hypothesize that exogenous geographic and climate factors that have beneficial effects on crop production, such as better soils, less overheating damage, more growing season precipitation and irrigation should make crop yield distributions less positively or more negatively skewed. We employ a large crop insurance dataset for corn, soybean, and wheat to find general support for the hypothesis. The novel empirical method optimally uses correlations between the first three moments and thus significantly improves estimation performance over existing methods.

Keywords: cross-moment correlation, generalized method of moments, von Liebig production technology

JEL Code: Q10, Q18, Q50.

1. Introduction

Concerns about how climate change could affect global food security, grain production, and corresponding crop market prices have been addressed by many, e.g., Lobell, Schlenker, and Costa-Roberts (2011). Yield mean responses to historical weather realizations, such as growing season temperature and precipitation, have been modeled explicitly (see, among others, Rosenzweig and Parry 1994). Results indicate statistically significant and large impacts on the yields of major crops including corn, soybeans, wheat and rice. Despite an extensive literature documenting the impacts of exogenous factors on crop yield means and variances, their effects on yield skewness are not well understood. It is an important oversight as a more negatively skewed distribution implies more frequent left-tail disasters. For U.S. crop insurance markets, there would be larger indemnity payouts and higher crop insurance subsidies. The objective of this study is to provide a rigorous theoretical and empirical understanding of the effects of exogenous geographic and climate-related factors, including soil quality, growing season temperature and precipitation, location, and irrigation, on the first three moments of crop yields. The estimates shed light on the role of geography in insurance rate setting. A carefully controlled geographic cross-sectional analysis of average climate variation also allows inferences on how temporal climate change in a given location would affect the first three moments of crop yields.

There have been recent developments in the understanding of the impact of weather conditions on crop yield distribution. For example, Lobell and Asner (2003) conclude that gradual temperature changes have reduced both corn and soybean yield by approximately 17% for one degree increase in growing season temperature over 1982-1998. Deschênes and Greenstone (2007) find beneficial effects of the increase in precipitation on corn and soybean yields while the increase in temperature is harmful. But the significance and magnitude of the effects vary across model specification. Nonlinear and asymmetric effects of temperature on crop yields are found by several studies including, e.g., Schlenker, Hanemann, and Fisher (2006) and Schlenker and Roberts (2009). In general, crop yield increases with temperature only up to a certain threshold, above which yield declines significantly. Based on the drought index constructed from monthly county-level rainfall and temperature, Yu and Babcock (2010) find that corn and soybeans are more drought tolerant in main Corn Belt states since 1980.

A related issue is how land quality shapes the moments of yield distributions. Although land quality is often measured with reference to impact on mean yield effects, effects on other moments are not so clear. The relation is important because land quality is spatially clustered and so its spatial distribution may explain some aspects of geographic regularities in crop disaster patterns. Although a voluminous literature exists on how soil quality affects yields (e.g., ISU 2006), the literature on higher moment effects is not well developed and we are not aware of any study on how soil quality affects yield skewness.

Another strand of literature studies the effects of various inputs on crop yield distribution using the stochastic production function (SPF) specification (Just and Pope 1978, 1979). In the SPF model, both mean and variance responses of crop yield to input factors are explicitly modeled and estimated. For example, McCarl, Villavicencio and Wu (2008) investigate the impact of historical climate change on the mean and variability of crop yield in the framework of SPF. Du, Hennessy and Yu (2012) extend the SPF model to accommodate crop yield skewness and apply it to assess how skewness responds to nitrogen use.

In the current study we propose an empirical model in the generalized method of moment (GMM) framework in which the impacts of exogenous input factors on the first three moments of crop yields are jointly estimated. The proposed GMM method is related to but different from the linear moment model (hereafter LMM) proposed in Antle (1983). The LMM method is flexible, easy to implement and has been widely applied in the literature. The applications include, for example, the effects of crop diversity on farm productivity and production risk (Di Falco and Chavas 2009), the role of production uncertainty and incomplete information in the adoption of irrigation technology (Koundouri, Nauges, and Tzouvelekas 2006), and asymmetric effects of inputs on distribution tails (Antle 2010). But in our Monte Carlo simulation the LMM method is found to be significantly biased for the coefficients of the skewness equation. After a modification of the third moment equation, the modified LMM method (M-LMM hereafter), largely corrects the biases. One distinctive feature of the proposed GMM estimation method is that we explicitly take into account the correlations between the first three moments. Compared with the LMM and M-LMM as well as a traditional OLS method using aggregated county level moment information, the GMM estimator outperforms the other methods, especially in small samples with fewer sampled counties.

The present study distinguishes itself from the existing studies by: (i) showing that commonly assumed production technologies lend support to the hypothesis that better growing conditions reduce yield skewness, (ii) making full use of a large insurance unit level yield dataset for corn, soybean and wheat, and (iii) optimally using correlations between the first three moments and thus significantly improving estimation performance over existing methods. Different from the majority of studies in the literature, which employ either aggregate county- or state-level data or relatively few farm-level observations (see a similar discussion in Claassen and Just 2010), this study utilizes a vast number of farm- or subfarm-level observations (over three million for each crop) and thus obtains sufficient power for the empirical study. Accounting for simultaneous correlation in a system of moment equations, which has been largely ignored in the literature, the proposed GMM method not only provides unbiased parameter estimates but also results in a substantial gain of statistical power.

Section 2 outlines the theoretical framework regarding how better growing conditions affect higher moments of the yield distribution. Section 3 explains the empirical method and data. Monte Carlo simulations are described and the results are discussed in Section 4. After the analysis of the estimation results, some concluding comments are offered.

2. Analytic Framework

We posit a standard stochastic production relation $y = Q(\varepsilon; z)$ where y is output, $z \in Z$ is some exogenous geographic factor and ε is a random variable that follows distribution $F(\varepsilon)$: $[\underline{\varepsilon}, \overline{\varepsilon}] \to [0,1]$. Here the geographic factor could be soil quality, precipitation or sunshine and it is assumed that ε is ordered such that $Q_\varepsilon(\cdot) > 0$, i.e., more is beneficial. Furthermore, $Q(\varepsilon; z)$ is held to be smoothly differentiable in that $\partial^2 \ln[Q_\varepsilon(\cdot)] / \partial\varepsilon\partial z$ exists on $(\varepsilon, z) \in [\underline{\varepsilon}, \overline{\varepsilon}] \times Z$. Our interest throughout is in yield skewness as commonly understood, i.e., $E[\{(y-\mu)/\sigma\}^3]$ where μ and σ are, respectively, the mean and standard deviation of yield.

Proposition 1. For any $z_1 > z_0$, the yield distribution of $y_1 = Q(\varepsilon; z_1)$ is more negatively (positively) skewed than that of $y_0 = Q(\varepsilon; z_0)$ if and only if $\partial \ln[Q_\varepsilon(\varepsilon; z)] / \partial\varepsilon$ is decreasing (increasing) in exogenous geographic factor z.

The proof is provided in the supplementary appendix. We provide three examples, showing different skewness implications for a shift in exogenous factor under three commonly used functional forms for modeling production relations.

EXAMPLE 1. Consider the logistic production technology where $y = (\varepsilon + z)/(1 + \varepsilon + z)$ and $\varepsilon + z > 0$. Then $\ln[Q_\varepsilon(\cdot)] = -2\ln(1 + \varepsilon + z)$, $d\ln[Q_\varepsilon(\cdot)] / d\varepsilon = -2/(1 + \varepsilon + z)$ and $d^2 \ln[Q_\varepsilon(\cdot)] / d\varepsilon dz = 2/(1 + \varepsilon + z)^2 > 0$. Thus an increase in the value of the geographic attribute increases skewness in this case.

EXAMPLE 2. For $y = 1 - e^{-\lambda(\varepsilon+z)}$ with $\lambda \in \overline{\mathbb{R}}_+$ then $\ln[Q_\varepsilon(\cdot)] = \ln(\lambda) - \lambda(\varepsilon + z)$, $d\ln[Q_\varepsilon(\cdot)]/d\varepsilon = -\lambda$ and $d^2\ln[Q_\varepsilon(\cdot)]/d\varepsilon dz = 0$. In this case an increase in the value of the geography attribute has no impact on skewness.

EXAMPLE 3. The von-Liebig law of the minimum production function is a widely accepted model of crop growth technology; see, e.g., Cerrato and Blackmer (1990). A standard representation of this production function is $y = \min[\tilde{y}, \theta]$ where $\tilde{y}$ is some latent production level that would occur absent a limiting factor, in this case represented by parameter θ. Limiting constraint θ might be of form $\theta = g(a)$ where $g(\cdot)$ is an increasing function and a is the state of seed genetics. In its most simple, linear, form the production function could be written as $y = \min[\varepsilon + z, \theta]$, where a point of non-differentiability (pnd) occurs at $\varepsilon + z = \theta$. Notice that random yield y can be positively skewed as ε can be positively skewed with support below $\theta - z$.

Because $y = \min[\varepsilon + z, \theta]$ is not an everywhere differentiable function we apply a smoothing kernel, i.e., a non-negative function $k(w): [-\delta, \delta] \to \overline{\mathbb{R}}_+$ so that the unsmoothed value is $\hat{y} = \min[w + \varepsilon + z, \theta]$. The intent of this example is to demonstrate that under reasonable smoothing over the linear von-Liebig function's pnd then an increase in the exogenous geographic factor leads to a negative skewing of the yield distribution.

As is standard, the kernel satisfies $\int_{-\delta}^{\delta} k(w)dw = 1$ and $k(-w) = k(w)$ where $\delta > 0$ (Parzen 1962). A rationale for kernel smoothing is randomness in input expression or spatial variability in soil endowments at the fine topography level (Berck and Helfand 1990). Condition

$\int_{-\delta}^{\delta} k(w)dw = 1$ is no more than a normalization constraint while $k(-w) = k(w)$ ensures symmetry around $\varepsilon + z$ so that any asymmetries are less likely to be imposed by the kernel.

For the uniform smoothing kernel the calculation $Q(\varepsilon; z, \theta) = \int_{-\delta}^{\delta} \min[w + \varepsilon + z, \theta] k(w)dw$ $= (2\delta)^{-1} \int_{-\delta}^{\delta} \min[w + \varepsilon + z, \theta]dw$ has three distinct regions of interest. The first is when the kernel smooths only below the pnd. In that case $\delta \leq \theta - \varepsilon - z$ and $Q(\varepsilon; z, \theta) =$ $(2\delta)^{-1} \int_{-\delta}^{\delta} \min[w + \varepsilon + z, \theta]dw = \varepsilon + z$ so that $Q_{\varepsilon}(\cdot) = 1$, $d\ln[Q_{\varepsilon}(\cdot)]/d\varepsilon = 0$ and $d^2 \ln[Q_{\varepsilon}(\cdot)]/d\varepsilon dz = 0$. Alternatively, the smoothing could occur entirely above the pnd. In that case the value of $d\ln[Q_{\varepsilon}(\cdot)]/d\varepsilon$ is not defined (being 0/0), but we will shortly show that its limiting value is negative infinity. The intermediate, and most interesting, case is where $0 \leq -\delta + \varepsilon + z < \theta < \delta + \varepsilon + z$. Then

$$
\begin{aligned}
Q(\varepsilon; z, \theta) &= (2\delta)^{-1} \int_{-\delta}^{\theta-\varepsilon-z} (w + \varepsilon + z)dw + (2\delta)^{-1} \int_{\theta-\varepsilon-z}^{\delta} \theta dw \\
&= \frac{\left[0.5w^2 + (\varepsilon + z)w\right]_{-\delta}^{\theta-\varepsilon-z}}{2\delta} + \frac{\theta(\delta - \theta + \varepsilon + z)}{2\delta} \\
&= \frac{(\theta + \varepsilon + z)\delta - 0.5\delta^2 - 0.5(\theta - \varepsilon - z)^2}{2\delta}.
\end{aligned}
\tag{1}
$$

So $Q_{\varepsilon}(\cdot) = (\delta + \theta - \varepsilon - z)/(2\delta)$, $d\ln[Q_{\varepsilon}(\cdot)]/d\varepsilon = -1/(\delta + \theta - \varepsilon - z)$ and $d^2 \ln[Q_{\varepsilon}(\cdot)]/d\varepsilon dz = -1/(\delta + \theta - \varepsilon - z)^2 < 0$. Thus, and by contrast with examples 1 and 2, an increase in the level of geographic factor makes skewness more negative. Figure 1 depicts the three cases. As the uniform kernel distributions shift to the right then relative curvature becomes more negative. The limiting value is where $z = \delta + \theta - \varepsilon$, i.e., the left-most tail of the smoothing kernel barely covers the pnd. Then the relative curvature expression $d\ln[Q_{\varepsilon}(\cdot)]/d\varepsilon$

has limiting value $\lim_{z\uparrow\delta+\theta-\varepsilon} d\ln[Q_\varepsilon(\cdot)]/d\varepsilon \to -\infty$.

We will show next that the uniform kernel is not at all special. Indeed symmetry is not needed either for yield under a smoothed linear von Liebig production function to become more negatively skewed upon an increase in an exogenous factor.

General Result for Smoothed Linear von Liebig Production

As above, the production technology is modeled as being of von-Liebig type with linear additive smoothing component $y = \min[w + \varepsilon + z, \theta]$, or

(2) $$y(w;\varepsilon + z) = \varepsilon + z + \min[w, \theta - \varepsilon - z],$$

where w follows distribution $K(w)$ with density $k(w)$. For future reference we write the survival function as $\bar{K}(\theta - \varepsilon - z) \equiv 1 - K(\theta - \varepsilon - z)$ and we note from an integration by parts that $\int_{-\delta}^{\delta} \min[w, \theta - \varepsilon - z] dK(w) = \theta - \varepsilon - z - \int_{-\delta}^{\theta-\varepsilon-z} K(w)dw$. Our interest is in mean, or aggregate, yield over the heterogeneity, i.e., [1]

(3) $$Q(\varepsilon; z) = \int_{-\delta}^{\delta} y(w;\varepsilon + z) dK(w) = \theta - \int_{-\delta}^{\theta-\varepsilon-z} K(w)dw.$$

Applying Proposition 1 to the kernelized production relation we obtain

(4) $$\frac{d^2 \ln[Q_\varepsilon(\varepsilon; z)]}{d\varepsilon dz} = \frac{K(w)k_w(w) - [k(w)]^2}{[K(w)]^2} = \frac{d^2 \ln[K(w)]}{dw^2}, \qquad w = \theta - \varepsilon - z;$$

where $K(w)$ is said to be logconcave (resp., logconvex) whenever $\ln[K(w)]$ is concave (resp., convex) in w. Thus we have

Proposition 2. For the kernel smoothed linear von Liebig production relation, an increase in

[1] Although Hennessy (2009) also considered a production function of type (2) above, there both ξ and η were random. By contrast with the production function in (3), the yield distribution function was bivariate. The skewness to be considered here is upon aggregating across realizations of ξ. As we will discuss, aggregation has important implications for skewness.

the exogenous geographic factor reduces (resp., increases) yield skewness whenever kernel distribution $K(w)$ is logconcave (resp., logconvex).

Notice that symmetry on $k(w)$ was not invoked in the analysis leading up to Proposition 2. We will show later that the logconcave distribution property bears no relation with symmetry of the form $k(w) = k(-w)$ so that one assumption has no implications for the plausibility of the other.

We will comment first, and briefly, on logconvexity of the distribution function as we believe it to be less plausible. In their extensive overview of the property and allied matters, Bagnoli and Bergstrom (2005) only identify one common distribution to possess a logconvex distribution. This is the mirror image Pareto distribution. To confirm why a logconvex distribution is implausible consider that it requires ratio $k(w)/K(w)$ to be increasing. The denominator, being a distribution function, is increasing so the density function must be increasing everywhere on its support. Even absent an assumption on kernel density symmetry, this is very implausible and is contrary to the concept of smoothing because the density would collapse to zero just beyond its mode $k(\delta)$.

Regarding logconcavity, this is true whenever $K(w)k_w(w) \leq [k(w)]^2$ and applies whenever density $k(w)$ is concave. More generally, it applies whenever the density function is log-concave, or $k(w)k_{ww}(w) \leq [k_w(w)]^2$ (An 1998). As such, it applies for the normal, uniform and beta distributions with interior mode (An 1998).[2] Notice that the beta distribution can be highly skewed. The logconcave density property and the unimodal property are strongly related (Ibragimov 1956), where a restricted form of unimodality (strong unimodality) implies

[2] The statistics literature on log-concave densities, distributions and survival functions is very large. See Dharmadhikari and Joag-dev (1988) for an extensive but dated review.

logconcave density. Kernel densities, and so kernel distributions, are typically assumed to be logconcave because they endow the smoothed function with analytically and intuitively appealing properties (Mukerjee 1988). The triweight kernel used elsewhere in this paper (Claassen and Just 2011) is logconcave, as is the Epanechnikov kernel that is optimal in minimizing density estimate mean square error (Epanechikov 1969) and under certain conditions when estimating regressions (Benedetti 1977).

3. Empirical Methods

In this section, we present the empirical GMM estimation method followed by a brief description of the LMM, M-LMM, and county-level OLS methods.

GMM method

Here we propose an empirical model in the generalized method of moment (GMM) framework that explicitly takes into account the possible cross moment correlations. Let y_{ij} be the crop yield of farmer *j* in county *i*, $i \in \{1,...,n\}$, $j \in \{1,...,n_i\}$, and let X_i be the $N \times K$ $(N = \sum_{i=1}^{n} n_i)$ vector of exogenous input factors considered for each county. We define the mean, variance and skewness of crop yield in an individual county, $g_1, g_2,$ and g_3, as the linear functions of covariates:

(5) $$\begin{cases} g_1(X_i) = \sum_{k=1}^{K} \alpha_k X_{ki} & \text{(mean)} \\ g_2(X_i) = \sum_{k=1}^{K} \beta_k X_{ki} & \text{(variance)} \\ g_3(X_i) = \sum_{k=1}^{K} \gamma_k X_{ki} & \text{(skewness)} \end{cases}$$

Notice that, by definition, variance is computed by use of an estimate for mean, while skewness is computed by use of estimates for mean and variance. That is, the fit of β parameters in the second moment equation depends on α parameters choices of the first moment equation, while the fit of γ parameters in the third moment equation depends on α and β parameters choices for both the first and the second moments. Incorporating cross moment equation information will improve the efficiency of parameter estimates. In the proposed GMM method, a system of mean, variance, and skewness equations is estimated jointly in the GMM framework to account for potential cross-moment correlations. The moment conditions that we will use to carry out GMM estimation are

$$(6) \qquad \begin{cases} E\left[y_{ij} - g_1(X_i)\right] \\ E\left[\left(y_{ij} - g_1(X_i)\right)^2 - g_2(X_i)\right] \\ E\left[\left(y_{ij} - g_1(X_i)\right)^3 / g_2(X_i)^{3/2} - g_3(X_i)\right] \end{cases}$$

The corresponding sample moment conditions are[3]

$$(7) \qquad \psi = n^{-1}\sum_{i=1}^{n} n_i^{-1}\sum_{j=1}^{n_i} \begin{bmatrix} y_{ij} - g_1(X_i) \\ \left(y_{ij} - g_1(X_i)\right)^2 / g_2(X_i) - 1 \\ \left(y_{ij} - g_1(X_i)\right)^3 / g_2(X_i)^{3/2} - g_3(X_i) \end{bmatrix}$$

Parameter vector $\Theta = \{\alpha_0, \alpha_1, ..., \alpha_k, \beta_0, \beta_1, ..., \beta_k, \gamma_0, \gamma_1, ..., \gamma_k\}$ can be estimated as $\widehat{\Theta} =$ $\arg\min \psi' \widehat{W} \psi$ where $\widehat{W}$ is a weighting matrix. The iterated GMM estimator is employed in our study.[4] In the first iteration, $\widehat{W}$ is assumed to be the identity matrix. But thereafter the iterated approach is applied where updated weights come from $\widehat{V}(\theta)$, the parameter vector variance-covariance matrix in the previous iteration. Specifically, the GMM estimator

[3] Note that the second moment is rescaled in order to speed up convergence.
[4] In the two GMM estimation methods, two-step feasible GMM applies the same procedure as the iterated GMM, in which the weighting matrix $\widehat{W}$ iteratively with $\widehat{\Theta}$ until the estimators satisfy pre-determined convergence criterion. Iterated GMM has slightly better finite-sample properties, although asymptotically the two methods are equivalent. See Hull (2005), Ch. 2 for a more detailed discussion.

proceeds as follows:

Step 0: Take $\widehat{W_i^0} = I_{3\times3}$, an identity matrix for an individual county i, $i \in \{1,...,n\}$, to be the diagonal elements of the initial weighting matrix $\widehat{W_0}$ (off-diagonal entries are all zero). Then minimize the objective function $\psi'\widehat{W_0}\psi$ to compute the GMM estimate $\widehat{\Theta}_0$, which is a consistent but not efficient estimator.

Step1: Take the GMM estimate to calculate $\widehat{W_i} = n_i^{-1}(n_i-1)^{-1}\sum_{j=1}^{n_i}(\hat{e}_{ij}-\bar{\hat{e}}_i)(\hat{e}_{ij}-\bar{\hat{e}}_i)'$, where

(8) $$\hat{e}_{ij} = \begin{bmatrix} y_{ij} - \sum_{k=1}^{K}\alpha_k X_{ki} / \sqrt{\sum_{k=1}^{K}\beta_k X_{ki}} \\ (y_{ij} - \sum_{k=1}^{K}\alpha_k X_{ki})^2 / \sum_{k=1}^{K}\beta_k X_{ki} - 1 \\ \left((y_{ij} - \sum_{k=1}^{K}\alpha_k X_{ki}) / \sqrt{\sum_{k=1}^{K}\beta_k X_{ki}} \right)^3 - \sum_{k=1}^{K}\gamma_k X_{ki} \end{bmatrix}'$$

and $\bar{\hat{e}}_i = n_i^{-1}\sum_{j=1}^{n_i}\hat{e}_{ij}$. The optimal matrix $\widehat{W}$ consists of the $\widehat{W_i}$'s as the diagonal elements.

Then minimize $\psi'\widehat{W}\psi$ to obtain $\widehat{\Theta}$, which should be asymptotically efficient.

Step 2. Iterate Step 1 until convergence, i.e., the iteration changes in the estimated parameters are less than a specified threshold.

One thing worth mentioning is that in the estimation of the optimal weighting matrix $\widehat{W}$ we take into account only the within county cross-moment correlation as in $\widehat{W_i}$, not the across county moment correlations. The reason is that we have insufficient information to match between farms in different counties for calculating variance-covariance matrix. This implies that the resulting optimal weighting matrix is a block diagonal matrix in which the diagonal entries are variance-covariance for individual counties, $\hat{W}_i$, and the entries outside the main diagonal are all zero.

The variance of our GMM estimator $\widehat{\Theta}$ is

(9) $\hat{V}(\widehat{\Theta}) = \left[\sum_{i=1}^{n} \Gamma_i'(\widehat{\Theta})\hat{W}_i\Gamma_i(\widehat{\Theta})\right]^{-1}$

where $\widehat{\Theta}$ is the iterated estimate of Θ, $\Gamma_i(\widehat{\Theta}) = [\partial m_i(\Theta)/\partial\Theta]|_{\theta=\hat{\theta}}$, the gradient matrix of $m_i(\theta)$, and $\hat{W}_i$ is the weighting matrix computed in Step 1 above.

LMM, M-LMM, and regional OLS methods

Given the notation in the GMM section above, the LMM method (Antle 1983) proceeds as follows:

(i) Estimate the mean regression model

(10) $y_{ij} = \sum_{k=1}^{K} \alpha_k X_{ki} + \delta_1$

to obtain the mean effect parameters $\hat{\alpha}$ and the associated error term $\hat{\delta}_1 = y_{ij} - \sum_{k=1}^{K} \alpha_k X_{ki}$

(ii) Estimate the variance and skewness parameters, $\hat{\beta}$ and $\hat{\gamma}$, from the regressions

(11−1) $(\hat{\delta}_1)^2 = \sum_{k=1}^{K} \beta_k X_{ki} + \delta_2$

(11−2) $(\hat{\delta}_1)^3 = \sum_{k=1}^{K} \gamma_k X_{ki} + \delta_3$

(iii) Deal with the heteroscedasticity issues in the estimation of eqns. (10) and (11) in two ways: (a) the Weighted Least Square (WLS) method (Antle 1983), and (b) the White heteroskedastic-consistent estimator for standard error estimation (Antle 2010). For eqn. (10), the weights for WLS are chosen as the fitted values in eqn. (11-1). But weights can be negative for some observations for eqn. (11). In this case, correction method (b) is feasible and easy to implement.

Eqn. (11-2) of the LMM is modified to correct for the estimation biases we find in the Monte Carlo simulation,[5]

(11−2') $(\hat{\delta}_1)^3 / \left(f_2(X_i, \hat{\beta})\right)^{3/2} = \sum_{k=1}^{K} \gamma_k X_{ki} + \delta_3$

[5] We will discuss the Monte Carlo simulation in a later section.

where $f_2(X_i, \hat{\beta})$ is the fitted value of eqn. (11-1). So in the M-LMM method, eqns. (10), (11-1) and $(11-2')$ are estimated using least square to obtain parameter estimates and we calculate White standard errors to account for heteroscedasticity.

A county-level OLS method is also used for model comparison and identifying initial values for the GMM method. First, using all yield observations in an individual county, we calculate sample mean, variance, and skewness. Then we regress the first three moments on county level explanatory variables. It is feasible in our case as our sample includes over 600 counties for each crop considered. It will be shown that compared with the proposed GMM method, the regional OLS method is restricted and disadvantaged with limited geographic coverage, i.e., fewer sampled counties.

4. Data

Now we turn to the discussion of crop yield data and county level explanatory variables considered in this study. Crop insurance unit level yield data for corn and soybean in 13 states (IL, IN, IA, KS, MI, MN, MO, NE, ND, OH, OK, SD, and WI) and wheat yield in 11 states (all states listed above except IA and WI) over 1990-2009 are obtained from the Risk Management Agency (RMA) of the U.S. Department of Agriculture (USDA). The data contain up to 10 years yield history (likely non-consecutive) on each included unit, which has been insured under the federal crop insurance program over the sample period. The insured unit can comprise one or more fields on a farm. From the dataset, we restrict our sample to the units with 10 years of yield record.

To investigate crop yield distribution, it is necessary to remove the systematic components of yield variation, which link to the characteristics of the production sites in both spatial and temporal dimensions (Claassen and Just 2010). Following Claassen and Just (2011), unit level crop yields are de-trended in the following three steps:[6]

(i) Estimate each county's trend yield, $\hat{y}^c$, using a nonparametric local regression method.

(ii) Transfer the county level trend yield to those of individual units, $\hat{y}^u$, in a multiplicative model, $\hat{y}^u = \phi^u \hat{y}^c$, where ϕ^u is the unit-specific productivity measure.

(iii) Obtaining the corresponding unit level random error $\varepsilon^u = y - \hat{y}^u$, where y is the unit level yield observation.

To ensure positive yield and full utilization of all the available data across the whole sample period, we transfer all de-trended historical yield to the yield of 2009 by adding back the county-level trend yield of 2009 and adjusting for unit-specific productivity.[7] Now we obtain the yield datasets with a large number of observations (over four million for each crop) that can be treated as being generated from the distribution of a particular year.

The exogenous input factors (X) considered to be influencing factors of crop yield distribution include a variety of geographic and climate related factors including soil quality, growing season precipitation and temperature, county location, and irrigation.[8] In what follows, we discuss the construction of these county-level independent variables and their expected effects on crop yield distribution.

Soil quality

County level average soil quality is represented by the proportion of land acreage under certain Land Capacity Classes (LCC) and sub-classes.[9] Land capability classes and subclasses are determined based on the soil's potential capacity to produce field crops or pasture. In the LCC classification, land belonging to Class I is the most productive with few land use limitations

[6] Interested readers are referred to eqns. (1), (2')-(4') in Claassen and Just (2010).
[7] This is essentially the reverse of the de-trending process, i.e., normalizing to a particular year.
[8] The impact of irrigation will be discussed at a later juncture.
[9] Readers are referred to the National Soil Survey Handbook for more information, which is available at http://soils.usda.gov/technical/handbook/contents/part622.html.

and Classes II-IV are land with moderate to severe limitations but are still suitable to cultivation. Land in Classes V-VIII is unsuitable for agricultural production. Thus for individual counties, we focus on the LCC classes I to IV.

We use the county level LCC data contained in the National Resource Inventory (NRI) dataset. For a given county the soil quality measure is constructed as the percentage of land acres classified as Classes I and II in the total acreage of Classes I-IV. Counties with high constructed soil quality measure represent good soil productivity are expected to have relatively high crop yield.

Growing degree days

Growing degree days (GDD) has been widely applied in the literature to measure the heat accumulation during a crop's growing season (see, e.g., Deschênes and Greenstone 2007; Schlenker, Hanemann and Fisher 2006; Schlenker and Roberts 2009). GDD is defined as the sum of degrees in the range between lower and upper thresholds over a specific time period. The temperature thresholds are 8°C and 32°C for corn and soybeans, and 0°C and 25°C for wheat. April-September is chosen as the growing season for corn and soybeans, April-August for wheat. County-level GDD is calculated as the annual average growing degree days in growing season over the period of 1975-2005. U.S. county-level monthly average growing degree days developed in Schlenker and Roberts (2009) are employed for our calculation.[10]

Overheating is found to be harmful to crop growth and yield production. Temperature above 34°C is considered to be overheated (Schlenker, Hanemann and Fisher 2006).[11] In this study, we construct the variable GDD34 to capture the effect of overheating on crop yield

[10] A detailed data description and original dataset can be downloaded at http://www.columbia.edu/~ws2162/dailyData/.

[11] Following the literature, the over-heat threshold is assumed to be the same for corn, soybeans, and wheat.

distribution. GDD34 is calculated as the county average growing degree days with temperature above 34°C over the period of 1975-2005.

Precipitation

The precipitation variable denoted by Prec is constructed to represent total amount of rainfall during each crops growing season. Similarly, it is averaged over the 31 year period for individual counties in the sample. The original dataset is also developed in Schlenker and Roberts (2009).

Location

The x and y coordinates of a county centroid (denoted by Location-X and Location-Y) are included to indicate the geographic location of each county. The county level spatial centroids are obtained from the US Census Bureau.[12] Inclusion of location is a judgement call as county geographic location tends to be strongly correlated with weather condition as counties located further east are likely associated with increasing precipitation and counties further north could be linked with decreasing growing degree days. On the other hand, we attempt to use the location variables to capture county level non-climate features such as distance to market, planting date choice, and infrastructure problems that aren't fully captured by the climate variables. So the two location variables are used to test for the robustness of the parameter estimates.

Irrigation is expected to have a significant impact on crop yield distributions. But it is well known that a farmer's irrigation decision is endogenous and is affected by a number of factors including, for example, surface and/or underground water quality and availability, climate, and growing condition. More importantly it could potentially weaken the effects of the included climate variables on crop yield. Therefore we first conduct the empirical analysis on the

[12] The county spatial data are available at ftp://ftp.igsb.uiowa.edu/gis_library/USA/us_counties.htm (accessed March 22, 2012).

irrigated and non-irrigated land separately and then combine the two data sets to investigate the effect of irrigation.

4. Monte Carlo Simulation

The simulation scheme is set up as follows. As discussed above, the mean, variance, and skewness functions, g_1, g_2, and g_3, of county i in eqn. (5), are defined as functions of county-level exogenous input factors of X_i. For simplification, we consider only two exogenous factors, which are randomly sampled from normal distribution, i.e., $X_1 \sim \text{Normal}(10,5)$ $X_2 \sim \text{Normal}(1,1)$. The corresponding coefficients for the first three moment equations including the constant term are $(\alpha_k, \beta_k, \gamma_k),\ k \in \{1,2,3\}$. We generate random samples for farm j of county i as the following:

(12) $$y_{ij} = g_1(X_i) + \sqrt{g_2(X_i)}\varepsilon_{ij},\ \ \varepsilon_{ij} \sim \text{Skewed Normal}(\xi, \omega, \eta).$$

where the random error ε_{ij} is drawn from a skewed normal distribution with location, scale, and shape parameters ξ, ω, and η, respectively.[13] Thus we have

(13) $$\begin{cases} E(\varepsilon_{ij}) = \xi + \omega\tau\sqrt{2/\pi},\ \ \tau = \eta/\sqrt{1+\eta^2} \\ V(\varepsilon_{ij}) = \omega^2(1 - 2\tau^2/\pi) \\ \text{skew}(\varepsilon_{ij}) = 0.5(4-\pi)(\tau\sqrt{2/\pi})^3/(1-2\tau^2/\pi)^{3/2} \end{cases}$$

We solve for (ξ, ω, η) for county i to satisfy the following conditions

(14) $$\begin{cases} E(\varepsilon_{ij}) = 0 \\ V(\varepsilon_{ij}) = 1 \\ \text{skew}(\varepsilon_{ij}) = g_3(X_i) \end{cases}$$

[13] Note that given true parameters in generated X_i's only those satisfying $-1 \le g_3(X_i) \le 1$ are selected, which is a restriction of the skewed normal distribution and also ensure the solvability of eqn. (14).

By doing so, we generate the error terms ε_{ij} satisfying

(15) $E(\varepsilon_{ij})=0$, $E(\varepsilon_{ij}^2)=1$, $E(\varepsilon_{ij}^3)\neq 0$, $E(\varepsilon_{ij}^4)\neq 0$, and $E(\varepsilon_{ij}^5)\neq 0$.

Therefore the off-diagonal elements of the covariance matrix of the first three moments, which are functions of the quantities defined in eqn. (15), are not necessarily zero.[14]

According to the simulation scheme described above, we generate random samples of y_{ij} for county i. The yield samples are then employed for the estimation of the LMM, M-LMM, OLS and GMM models. To compare the estimation results across models, three statistics are calculated including Monte Carlo (MC) relative biases, MC variance, and mean squared error (MSE). They are defined as:

(a) Monte Carlo (MC) relative biases: $M^{-1}\sum_{m=1}^{M}(\hat{\Theta}_m-\Theta_{true})$, where M is the number of simulations, $M=1000$ in our case; Θ_{true} is the true parameters for generating the random samples, and $\hat{\Theta}_m$ is the estimated parameters from the *m*th random sample.

(b) MC variances: $\mathrm{var}(\hat{\Theta})$, $\hat{\Theta}=(\hat{\Theta}_1,\hat{\Theta}_2,...,\hat{\Theta}_M)'$.

(c) Mean squared error: $M^{-1}\sum_{m=1}^{M}(\hat{\Theta}_m-\Theta_{true})^2$.

Table 1 presents the simulation estimation and model comparison results for 10 counties $(n=10)$ and 200 observations for each county $(n_i=200)$. The results for 100 counties $(n=100)$ and 200 observations for each county $(n_i=200)$ are included in the Supplemental Materials. The results indicate that for the coefficients in the skewness equation, γ_0, γ_1, and γ_2, when compared with other methods, LMM generates significant biases. After normalization using fitted values of the second moment equation, The M-LMM method largely corrects the biases. The OLS and GMM methods further reduce the biases. For the mean and

[14] The derivation of the variance covariance matrix is provided in the Supplemental Materials.

variance equations, LMM and M-LMM generate the same results as they are essentially the same method. The OLS and GMM methods generate slightly better results for some coefficients. The MC biases are comparable between the OLS and GMM methods, while OLS performs better in the large samples with more sampled counties.

For the MC variances, after taking into account the cross-moment correlation the GMM method outperforms all other methods, especially LMM, on all coefficients. In the large samples, OLS generates comparable results in the skewness coefficients. It is not surprising given that the OLS method is based on aggregated county level moment information. We observe significant MSE reduction of the GMM method on all coefficients in the small sample over all other methods, especially the LMM on the coefficients of the skewness equation. Due to large biases, the LMM method has the largest MSE on the skewness parameters. M-LMM shows improvement over LMM after modifying the third moment equation, even in the large sample.

5. Analysis of Results

Table 2 presents the GMM estimation results for corn, soybeans, and wheat on non-irrigated land.[15] For each crop, column I in Table 2 includes the estimated impact on crop yield moments of geographic and climate-related variables, while column II includes two location variables, Location-X and Location-Y. The purpose of doing so is to distinguish the effect of county location from those of other included exogenous factors and to test for robustness of main results. We don't see consistent effects of county geographic location on the first three moments of yield distribution across crops, although all estimated coefficients are statistically significant at 1% level. As expected, soil quality (LCC) and growing season precipitation

[15] To save space, not all estimation results of LMM, M-LMM, and OLS methods are presented. The complete set of estimates is available upon request.

(Prec) are found to have significant and positive effects on the mean yield across crops and model specifications. Overheating temperature during growing season (GDD34) damages the mean yield of corn, soybean, and wheat. In the variance equation, the results indicate that good-quality soils (LCC) tend to associate with relatively high yield variation. For the third moment of the crop yield distribution, better average soil quality in an individual county represented by higher LCC makes unit-level crop yields of corn, soybeans and wheat more negatively skewed. During the growing season more overheating days tend to make the crop yield distribution more positively or less negatively skewed. The estimates confirm the hypotheses described in the Propositions 1 and 2.

GMM estimation results for crop yield on irrigated land are reported in Table 3. Across the three crops and model specifications, we don't find consistent impact of all included variables on the first three moments of yield distribution. Especially for soybean and wheat, the majority of the estimated coefficients are not significant. This is not surprising given the fact that the crop yield data on irrigated land are very limited, which include only 53 counties for corn, 17 for soybeans and 7 for wheat.

To separate the impact of irrigation on crop yield distribution, we combine crop yield data for non-irrigated and irrigated land and report the GMM estimation results in Table 4. Irrigation is found to significantly increase the mean yield and doesn't have a consistent impact on crop yield variation. For the skewness, after separating the irrigation effect, growing season precipitation (Prec) decreases crop yield skewness. Also the results show that irrigation makes corn and soybean, but not wheat, yield distributions more negatively skewed.

It is readily surmised that the irrigation decision is endogenous and is largely determined by hydrological and institutional factors including surface and underground water availability, water right and government policies (see, e.g., Moore and Negri 1992, Schlenker, Hanemann and Fisher 2005). Therefore the estimated irrigation effect could be potentially biased. Finding appropriate instrumental variables is challenging and will be addressed in future study. Table 5 summarizes the results in regard to Proposition 2.

6. Conclusion

We hypothesize that crop yield distributions will become more negatively skewed when growing conditions become more benign. Making full use of a large farm level yield dataset for corn, soybean, and wheat, we quantify the impact of exogenous geographic and climate related factors including soil quality, growing season temperature and precipitation, location, and irrigation on the first three moments of yield distribution. In general, exogenous geographic and climate related factors having beneficial effects on crop production, such as better soils, less overheating damage, and more precipitation, are found to support our hypothesis. The novel empirical method optimally uses correlations between the first three moments and thus significantly improves estimation performance over existing methods.

References

An, M.-Y. 1998. "Logconcavity versus Logconvexity: A Complete Characterization." *Journal of Economic Theory* 80(2):350–369.

Antle, J.M. 1983. "Testing the Stochastic Structure of Production: A Flexible Moment-Based Approach." *Journal of Business & Economic Statistics* 1(3):192-201.

Antle, J.M. 2010. "Asymmetry, Partial Moments and Production Risk. *American Journal of Agricultural Economics* 92(5): 1294-1309.

Bagnoli, M., and T.C. Bergstrom. 2005. "Log-Concave Probability and its Applications." *Economic Theory* 26(2):445–469.

Benedetti, J.K. 1977. "On the Nonparametric Estimation of Regression Functions." *Journal of the Royal Statistical Society. Series B (Methodological)* 39(2):248–253.

Berck, P., and G. Helfand. 1990. "Reconciling the von Liebig and Differentiable Crop Production Functions." *American Journal of Agricultural Economics* 72(4):985–996.

Cerrato, M.E., and A.M. Blackmer. 1990. "Comparison of Models for Describing Corn Yield

Response to Nitrogen Fertilizer." *Agronomy Journal* 82(January/February):138–143.

Claassen, R., and R.E. Just. 2011. "Heterogeneity and Distributional Form of Farm Yields." *American Journal of Agricultural Economics* 93(1):144–160.

Deschênes, O. and Greenstone, M. 2007. "The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather." *American Economic Review* 97(1): 354-385.

Dharmadhikari, S., and K. Joag-dev. 1988. *Unimodality, Convexity, and Applications*. Academic Press: Boston.

Di Falco, S., and J.P. Chavas. 2009. "On Crop Biodiversity, Risk Exposure, and Food Security in the Highlands of Ethiopia." *American Journal of Agricultural Economics* 91(3):599-611.

Du, X., D.A. Hennessy, and C.L. Yu. 2012. "Testing Day's Conjecture that More Nitrogen Decreases Crop Yield Skewness." *American Journal of Agricultural Economics* 94(1): 225-237.

Epanechnikov, V.A. 1969. "Nonparametric Estimates of a Multivariate Probability Density." *Theory of Probability and its Applications* 14(1):153–158.

Hennessy, D.A. 2009. "Crop Yield Skewness under Law of the Minimum Technology." *American Journal of Agricultural Economics* 91(1):197–208.

Hull, A.R. 2005. *Generalized Method of Moments*. Oxford, NY: Oxford University Press.

Ibragimov, I.A. 1956. "On the Composition of Unimodal Distributions." *Theory of Probability and its Applications* 1(2):255–266.

Iowa State University Extension (ISU). "Corn Suitability Ratings - An Index to Soil Productivity." Extension Publication PM1168, 2006.

Just, R.E., and R.D. Pope. 1978. "Stochastic Specification of Production Function and Economic Implication." *Journal of Econometrics* 7(1): 67-86.

--. 1979. "Production Function Estimation and Related Risk Considerations." *American Journal of Agricultural Economics* 81(2): 287-304.

Koundouri, P., C. Nauges, and V. Tzouvelekas. 2006. "Technology Adoption under Production Uncertainty: Theory and Application to Irrigation Technology." *American Journal of Agricultural Economics* 88(3):657–670.

Lobell, D.B., and G.P. Asner. 2003. "Climate and Management Contributions to Recent Trends in U.S. Agricultural Yields." *Science* 299: 1032.

McCarl, B.A., X. Villavicencio, and X. Wu. 2008. "Climate Change and Future Analysis: Is Stationarity Dying?" *American Journal of Agricultural Economics* 90(5): 1241-1247.

Moore, M.R., and D.H. Negri. "A Multicrop Production Model of Irrigated Agriculture, Applied to Water Allocation Policy of the Bureau of Reclamation." *Journal of Agricultural and Resource Economics* 17(1): 29-43.

Mukerjee, H. 1988. "Monotone Nonparametric Regression." *Annals of Statistics* 16(1):741–750.

Parzen, E. 1962. "On Estimation of a Probability Density Function and Mode." *Annals of Mathematical Statistics* 33(3):1065–1076.

Ramirez, O.A., S. Misra, and J. Field. 2003. "Crop Yield Distributions Revisited." *American Journal of Agricultural Economics* 85(1):108–120.

Rao, M.M. 1984. *Probability Theory and Applications*. Academic Press: Orlando.

Rosenzweig, C., and M.L. Parry. 1994. "Potential Impact of Climate Change on World Food Supply." *Nature* 367(Jan. 13th): 133-138.

Schlenker, W., M.W. Hanemann, and A.C. Fisher. 2005. "Will U.S. Agriculture Really Benefit from Global Warming? Accounting for Irrigation in the Hedonic Approach." *American Economic Review* 95(1): 395-406.

--. 2006. The Impact of Global Warming on U.S. Agriculture: An Econometric Analysis of Optimal Growing Conditions. *Review of Economics and Statistics* 88(1): 113-125.

Schlenker, W. and M.J. Roberts. 2009. Nonlinear Temperature Effects Indicate Severe Damages to U.S. Crop Yields under Climate Change. *Proceedings of the National Academy of Sciences* 106(37): 15594-15598.

Yu, T., and B.A. Babcock. 2010. "Are U.S. Corn and Soybeans Becoming More Drought Tolerant?" *American Journal of Agricultural Economics* 92(5):1310–1323.

Zwet van, W.R. *Convex Transformations of Random Variables*. Amsterdam: Matematisch Centrum, 1964.

Table 1. Simulation results for 10 counties and 200 observations for each county

Coefficient	LMM (I)	M-LMM (II)	OLS (III)	GMM (IV)	IV/I	IV/II	IV/III
MC relative biases ($\times 10^{-3}$)							
Mean- α_0	13.6028	13.6028	13.6028	9.1836	0.6750	0.6750	0.6750
α_1	0.5460	0.5460	0.5460	0.1520	0.2780	0.2780	0.2780
α_2	2.0243	2.0243	2.0243	1.1470	0.5670	0.5670	0.5670
Variance- β_0	19.5496	19.5496	19.6725	-2.9777	-0.1520	-0.1520	-0.1510
β_1	2.5518	2.5518	3.0829	-2.9629	-1.1610	-1.1610	-0.9610
β_2	5.1001	5.1001	6.8713	-5.2623	-1.0320	-1.0320	-0.7660
Skewness- γ_0	-91.9585	73.5376	31.0560	-9.6676	0.1050	-0.1310	-0.3110
γ_1	280.2923	-12.1866	-6.3189	-10.1452	-0.0360	0.8320	1.6060
γ_2	518.5034	-33.2444	-16.7233	-23.4231	-0.0450	0.7050	1.4010
MC variances ($\times 10^{-3}$)							
Mean- α_0	0.1924	0.1924	0.1924	0.1674	0.8700	0.8700	0.8700
α_1	0.0120	0.0120	0.0120	0.0093	0.7740	0.7740	0.7740
α_2	0.2417	0.2417	0.2417	0.2232	0.9230	0.9230	0.9230
Variance- β_0	0.6862	0.6862	0.6918	0.4625	0.6740	0.6740	0.6680
β_1	0.0519	0.0519	0.0522	0.0263	0.5070	0.5070	0.5040
β_2	1.0144	1.0144	1.0253	0.7654	0.7550	0.7550	0.7470
Skewness- γ_0	6.4216	3.0120	1.2613	1.1809	0.1840	0.3920	0.9360
γ_1	0.6340	0.1649	0.0767	0.0677	0.1070	0.4110	0.8840
γ_2	9.9877	1.7260	1.2310	1.1051	0.1110	0.6400	0.8980
MSE ($\times 10^{-3}$)							
Mean- α_0	2.1051	2.1051	2.1051	1.7546	0.8340	0.8340	0.8340
α_1	0.1197	0.1197	0.1197	0.0924	0.7720	0.7720	0.7720
α_2	2.4167	2.4167	2.4167	2.2293	0.9220	0.9220	0.9220

Variance- β_0	7.2302	7.2302	7.2916	4.6246	0.6400	0.6400	0.6340
β_1	0.5245	0.5245	0.5300	0.2712	0.5170	0.5170	0.5120
β_2	10.1499	10.1499	10.2796	7.6663	0.7550	0.7550	0.7460
Skewness- γ_0	72.5440	65.4679	13.5520	11.8788	0.1640	0.3350	0.8770
γ_1	84.8914	1.7946	0.8050	0.7790	0.0090	0.4340	0.9680
γ_2	368.5235	18.3328	12.5649	11.5779	0.0310	0.6320	0.9210

Table 2. GMM Estimation Results of Corn, Soybeans, and Wheat Yield Moments on Non-irrigated Land (Standard errors are in the parentheses)

	Corn		Soybeans		Wheat	
	I	II	I	II	I	II
Mean						
Constant	1.13[c]	0.27[c]	0.08[c]	0.54[c]	0.12[c]	1.95[c]
	(0.003)	(0.009)	(0.0007)	(0.003)	(0.003)	(0.006)
LCC	0.04[c]	0.21[c]	0.04[c]	0.06[c]	0.20[c]	0.13[c]
	(0.001)	(0.001)	(0.0003)	(0.0003)	(0.0005)	(0.0005)
GDD	0.17[c]	0.43[c]	0.07[c]	-0.07[c]	0.001	-0.41[c]
	(0.001)	(0.004)	(0.0004)	(0.001)	(0.001)	(0.003)
GDD34	-0.38[c]	-0.33[c]	-0.05[c]	-0.09[c]	-0.03[c]	-0.01[c]
	(0.001)	(0.001)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Prec	0.36[c]	0.47[c]	0.21[c]	0.13[c]	0.26[c]	-0.16[c]
	(0.002)	(0.002)	(0.0005)	(0.0007)	(0.002)	(0.002)
Location-X		-0.20[c]		-0.10[c]		0.47[c]
		(0.0009)		(0.0003)		(0.0009)
Location-Y		0.08[c]		-0.16[c]		-0.16[c]
		(0.002)		(0.0008)		(0.0006)
Variance						
Constant	-0.04[c]	-0.19[c]	0.03[c]	-0.01[c]	-0.01[c]	0.08[c]
	(0.001)	(0.0009)	(0.00004)	(0.0002)	(0.0001)	(0.0004)
LCC	0.01[c]	0.01[c]	0.003[c]	0.002[c]	0.004[c]	0.004[c]
	(0.0001)	(0.0001)	(0.00001)	(0.00002)	(0.00002)	(0.00002)
GDD	0.08[c]	0.15[c]	-0.01[c]	0.004[c]	0.0004[c]	-0.03[c]
	(0.0002)	(0.0003)	(0.00001)	(0.0001)	(0.0001)	(0.0002)
GDD34	0.10[c]	-0.01[c]	0.005[c]	0.005[c]	0.002[c]	0.002[c]
	(0.0003)	(0.0002)	(0.00001)	(0.00001)	(0.00001)	(0.00001)
Prec	-0.07[c]	-0.01[c]	-0.002[c]	0.002[c]	0.03[c]	0.02[c]
	(0.0004)	(0.0002)	(0.00003)	(0.00005)	(0.0001)	(0.00016)
Location-X		-0.004[c]		0.004[c]		-0.01[c]
		(0.0001)		(0.00002)		(0.00003)
Location-Y		-0.02[c]		0.01[c]		-0.01[c]
		(0.0002)		(0.0001)		(0.00005)
Skewness						
Constant	1.05[c]	8.28[c]	-2.89[c]	1.88[c]	-2.67[c]	-1.56[c]
	(0.04)	(0.15)	(0.04)	(0.14)	(0.10)	(0.17)
LCC	-0.14[c]	-0.33[c]	-0.48[c]	-0.02	-0.98[c]	-0.39[c]
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
GDD	-1.60[c]	-2.18[c]	1.38[c]	-0.73[c]	1.03[c]	-0.90[c]
	(0.03)	(0.06)	(0.03)	(0.06)	(0.05)	(0.06)
GDD34	0.52[c]	1.81[c]	0.06[c]	0.90[c]	0.12[c]	0.15[c]
	(0.007)	(0.01)	(0.01)	(0.01)	(0.004)	(0.004)
Prec	1.31[c]	-4.50[c]	-0.19[c]	-0.98[c]	-0.80[c]	2.88[c]
	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)	(0.07)
Location-X		0.92[c]		0.57[c]		-2.03[c]
		(0.02)		(0.01)		(0.05)
Location-Y		0.23[c]		-0.57[c]		0.06[c]
		(0.04)		(0.03)		(0.01)

Note: a, b and c denote significance at 0.10, 0.05, and 0.01 levels, respectively.

Table 3. GMM Estimation Results of Corn, Soybeans, and Wheat Yield Moments on Irrigated Land (Standard errors are in the parentheses)

	Corn		Soybeans		Wheat	
	I	II	I	II	I	II
Mean						
Constant	2.25^{c}	5.60^{c}	0.55^{c}	-0.05	0.23^{b}	-3.64
	(0.02)	(0.07)	(0.02)	(0.10)	(0.11)	(7.20)
LCC	0.28^{c}	0.36^{c}	-0.002	0.01^{b}	0.08	0.13
	(0.003)	(0.004)	(0.003)	(0.005)	(0.05)	(0.23)
GDD	-0.64^{c}	-1.20^{c}	0.16^{c}	0.35^{c}	0.09	1.58
	(0.02)	(0.02)	(0.02)	(0.03)	(0.11)	(2.73)
GDD34	0.07^{c}	0.09^{c}	0.02^{c}	0.04^{c}	-0.02	-0.10
	(0.002)	(0.002)	(0.002)	(0.004)	(0.03)	(0.19)
Prec	0.86^{c}	-0.42^{c}	-0.30^{c}	-0.28^{c}	-0.005	-0.54
	(0.001)	(0.03)	(0.01)	(0.04)	(0.33)	(1.29)
Location-X		1.57^{c}		-0.19^{c}		-0.42
		(0.03)		(0.04)		(1.12)
Location-Y		-0.30^{c}		0.17^{c}		0.66
		(0.009)		(0.01)		(1.10)
Variance						
Constant	-0.44^{c}	1.23^{c}	0.006^{c}	-0.11^{c}	-0.02	1.94
	(0.007)	(0.03)	(0.004)	(0.02)	(0.03)	(1.73)
LCC	-0.08^{c}	-0.09^{c}	0.004^{c}	-0.001	-0.05^{c}	0.02
	(0.002)	(0.002)	(0.001)	(0.001)	(0.02)	(0.06)
GDD	0.29^{c}	0.17^{c}	0.01^{c}	0.03^{c}	-0.04	-0.81
	(0.006)	(0.008)	(0.003)	(0.006)	(0.03)	(0.66)
GDD34	0.05^{c}	0.01^{c}	-0.0004	-0.004^{c}	0.01^{a}	0.02
	(0.001)	(0.001)	(0.0005)	(0.001)	(0.007)	(0.05)
Prec	-0.07^{c}	-0.93^{c}	-0.02^{c}	0.02^{c}	0.25^{c}	0.34
	(0.005)	(0.02)	(0.003)	(0.008)	(0.09)	(0.31)
Location-X		-0.75^{c}		-0.05^{c}		-0.03
		(0.01)		(0.008)		(0.27)
Location-Y		-0.26^{c}		0.01^{c}		-0.44^{a}
		(0.005)		(0.003)		(0.26)
Skewness						
Constant	1.42^{c}	4.48^{c}	0.84	2.37	-6.31^{c}	-32.95
	(0.27)	(0.81)	(1.24)	(5.37)	(0.98)	(66.45)
LCC	-0.26^{c}	-0.71^{c}	-0.03	-0.17	1.14^{b}	2.84
	(0.03)	(0.03)	(0.15)	(0.30)	(0.52)	(2.29)
GDD	-2.02^{c}	-4.53^{c}	-1.99^{b}	-2.36^{a}	4.71^{c}	13.84
	(0.20)	(0.31)	(0.92)	(1.41)	(1.10)	(25.71)
GDD34	0.33^{c}	0.41^{c}	0.41^{c}	0.23	-0.71^{c}	-1.64
	(0.03)	(0.03)	(0.14)	(0.25)	(0.23)	(1.85)
Prec	1.13^{c}	3.63^{c}	1.93^{b}	2.09	-8.95^{c}	13.35
	(0.14)	(0.28)	(0.89)	(2.21)	(3.14)	(13.35)
Location-X		-0.64^{c}		0.70		-7.70
		(0.31)		(2.21)		(10.02)
Location-Y		-1.00^{c}		-0.83		2.73
		(0.10)		(0.68)		(10.19)

Note: a, b and c denote significance at 0.10, 0.05, and 0.01 levels, respectively.

Table 4. GMM Estimation Results of Corn, Soybeans, and Wheat Yield Moments on Non-irrigated and Irrigated Land (Standard errors are in the parentheses)

	Corn		Soybeans		Wheat	
	I	II	I	II	I	II
Mean						
Constant	1.10^{c}	0.64^{c}	0.08^{c}	0.54^{c}	0.13^{c}	1.84^{c}
	(0.002)	(0.01)	(0.0007)	(0.003)	(0.003)	(0.006)
LCC	0.18^{c}	0.21^{c}	0.04^{c}	0.06^{b}	0.19^{c}	0.13^{c}
	(0.0008)	(0.0009)	(0.0003)	(0.0003)	(0.0005)	(0.0005)
GDD	0.03^{c}	0.21^{c}	0.07^{c}	-0.07^{c}	-0.0004	-0.37^{c}
	(0.001)	(0.003)	(0.0004)	(0.001)	(0.001)	(0.003)
GDD34	-0.16^{c}	-0.20^{c}	-0.05^{c}	-0.09^{c}	-0.03^{c}	-0.01^{c}
	(0.0006)	(0.0007)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Prec	0.46^{c}	0.51^{c}	0.21^{c}	0.13^{c}	0.26^{c}	-0.18^{c}
	(0.002)	(0.002)	(0.0005)	(0.0007)	(0.001)	(0.002)
Location-X		-0.14^{c}		-0.10^{c}		0.47^{c}
		(0.0008)		(0.0003)		(0.0009)
Location-Y		0.04^{c}		-0.16^{c}		-0.16^{c}
		(0.002)		(0.0007)		(0.0006)
Irrigation	0.46^{c}	0.44^{c}	0.17^{c}	0.13^{c}	0.19^{c}	0.14^{c}
	(0.0006)	(0.0008)	(0.0004)	(0.0004)	(0.001)	(0.001)
Variance						
Constant	-0.16^{c}	-0.11^{c}	0.03^{c}	-0.01^{c}	-0.01^{c}	0.08^{c}
	(0.0003)	(0.0004)	(0.00004)	(0.0002)	(0.0001)	(0.0003)
LCC	0.01^{c}	0.02^{c}	0.004^{c}	0.002	0.004^{c}	0.004^{c}
	(0.00006)	(0.0001)	(0.00001)	(0.00002)	(0.00002)	(0.00002)
GDD	0.12^{c}	0.12^{c}	-0.01^{c}	0.004^{c}	-0.0009^{c}	-0.03^{c}
	(0.0002)	(0.0002)	(0.00001)	(0.00009)	(0.00007)	(0.0001)
GDD34	0.05^{c}	0.003^{c}	0.005^{c}	0.005^{c}	0.002^{c}	0.002^{c}
	(0.0002)	(0.0001)	(0.000001)	(0.00001)	(0.00001)	(0.00001)
Prec	-0.02^{c}	-0.04^{c}	-0.002^{c}	0.001^{c}	0.03^{c}	0.02^{c}
	(0.0001)	(0.0001)	(0.00003)	(0.00005)	(0.00006)	(0.0002)
Location-X		-0.005^{c}		0.004^{c}		-0.009^{c}
		(0.0001)		(0.00002)		(0.00003)
Location-Y		-0.04^{c}		0.01^{c}		-0.01^{c}
		(0.0001)		(0.00005)		(0.00005)
Irrigation	-0.03^{c}	-0.04^{c}	0.0007^{c}	0.003^{c}	0.03^{c}	0.03^{c}
	(0.0002)	(0.0002)	(0.00006)	(0.00006)	(0.0003)	(0.0003)
Skewness						
Constant	1.81^{c}	0.36^{c}	-3.20^{c}	1.63^{c}	-3.55^{c}	-2.07^{c}
	(0.04)	(0.13)	(0.04)	(0.14)	(0.10)	(0.17)
LCC	-0.63^{c}	-0.93^{c}	-0.38^{c}	-0.05^{c}	-0.98^{c}	-0.39^{c}
	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
GDD	-0.64^{c}	1.00^{c}	1.54^{c}	-0.59^{c}	1.35^{c}	-1.00^{c}
	(0.02)	(0.05)	(0.03)	(0.06)	(0.05)	(0.06)
GDD34	0.10^{c}	-0.07^{c}	0.02^{b}	0.85^{c}	0.11^{c}	0.16^{c}
	(0.005)	(0.008)	(0.01)	(0.01)	(0.004)	(0.004)
Prec	-0.72^{c}	-2.58^{c}	-0.21^{c}	-0.98^{c}	-0.86^{c}	3.44^{c}
	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)	(0.07)
Location-X		0.05^{c}		0.57^{c}		-2.44^{c}
		(0.02)		(0.01)		(0.05)
Location-Y		1.00^{c}		-0.55^{c}		0.10^{c}
		(0.03)		(0.03)		(0.01)

Irrigation	-1.05^{c}	-1.27^{c}	-0.59^{c}	-0.73^{c}	0.05^{c}	0.19^{c}
	(0.01)	(0.02)	(0.03)	(0.02)	(0.01)	(0.01)

Note: a, b and c denote significance at 0.10, 0.05, and 0.01 levels, respectively.

Table 5. Effects of Geographic Factors on Yield Skewness; 'Y' if as hypothesized and significant at 1% level, 'N' otherwise

	Corn		Soybeans		Wheat	
	I	II	I	II	I	II
Table 2						
LCC	Y	Y	Y	Y	Y	Y
GDD	Y	Y	N	Y	N	Y
GDD34	Y	Y	Y	Y	Y	Y
Precipitation	N	Y	Y	Y	Y	N
Table 3						
LCC	Y	Y	N	N	N	N
GDD	Y	Y	N	N	N	N
GDD34	Y	Y	Y	N	N	N
Table 4						
LCC	Y	Y	Y	Y	Y	Y
GDD	Y	N	N	Y	N	Y
GDD34	Y	N	N	Y	Y	Y
Precipitation	Y	Y	Y	Y	Y	N
Irrigation	Y	Y	Y	Y	N	N

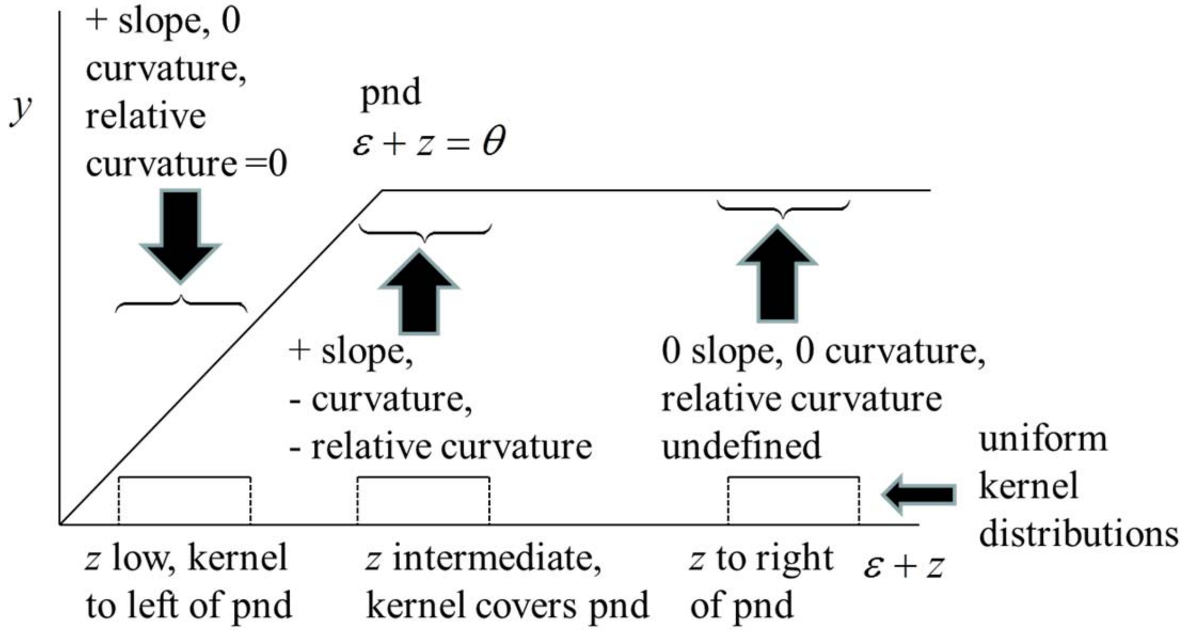

Figure 1. Effect of smoothing kernel location on relative curvature of smoothed production function.

Appendix A: Proof of Proposition 1

It is known from van Zwet (1964) that an increasing and concave transformation of a random variable makes it more negatively skewed. We will use this observation to compare the distributions of y_0 and y_1. Let ε^p solve $p = F(\varepsilon^p)$, i.e., it is the *p*th quantile. As $Q_\varepsilon(\varepsilon; z) > 0$ it follows that the *p*th quantile for y_0 is given by $y_0^p = Q(\varepsilon^p; z_0)$ so that inversion in the first argument provides $\varepsilon^p = Q^{-1}(y_0^p; z_0)$. We can perform the same quantile inversion for $y_1^p =$ $Q(\varepsilon^p; z_1)$ so that $\varepsilon^p = Q^{-1}(y_1^p; z_0) = Q^{-1}(y_0^p; z_0)$ and $y_1^p = Q(Q^{-1}(y_0^p; z_0); z_1)$. Because $Q_\varepsilon(\varepsilon; z) > 0$ ensures that yield quantiles follow that of ε, we may drop quantile notation and just write the transformation of random variables as $y_1 = Q(Q^{-1}(y_0; z_0); z_1) = J(y_0; z_0, z_1)$

We seek to identify conditions under which $J(y_0; z_0, z_1)$ is concave in its first argument. Write $\varepsilon = Q^{-1}(y_0; z_0)$ and $u = Q(\varepsilon; z_1)$ so that $du / dy_0 = Q_\varepsilon(\varepsilon; z_1)(d\varepsilon / dy_0)$. Now $d\varepsilon / dy_0 =$ $1/(dy_0 / d\varepsilon) = 1/Q_\varepsilon(\varepsilon; z_0)$ and therefore

$$\text{(A.1)} \quad \frac{du}{dy_0} = \frac{Q_\varepsilon(\varepsilon; z_1)}{Q_\varepsilon(\varepsilon; z_0)} > 0.$$

In addition,

$$\text{(A.2)} \quad \begin{aligned} \frac{d^2 u}{d(y_0)^2} &= \frac{1}{[Q_\varepsilon(\varepsilon; z_0)]^2} \left\{ Q_\varepsilon(\varepsilon; z_0) Q_{\varepsilon\varepsilon}(\varepsilon; z_1) \frac{d\varepsilon}{dy_0} - Q_\varepsilon(\varepsilon; z_1) Q_{\varepsilon\varepsilon}(\varepsilon; z_0) \frac{d\varepsilon}{dy_0} \right\} \\ &= \frac{Q_\varepsilon(\varepsilon; z_1)}{[Q_\varepsilon(\varepsilon; z_0)]^2} \left\{ \frac{d \ln[Q_\varepsilon(\varepsilon; z_1)]}{d\varepsilon} - \frac{d \ln[Q_\varepsilon(\varepsilon; z_0)]}{d\varepsilon} \right\}. \end{aligned}$$

As the term outside the parentheses is strictly positive, the curvature of $J(y_0; z_0, z_1)$ in y_0 is determined by the sign of the term in brackets, i.e., by the monotonicity of $d \ln[Q_\varepsilon(\varepsilon; z)] / d\varepsilon$ in z. The proposition follows from (A.1) and (A.2).

Appendix B: Derivation of the variance covariance matrix for the simulation studies

Given the simulation setup in Eqn. (12), we have the three moments equations for county i as

(B-1) $$\begin{cases} m_{i1} = n_i^{-1}\sum_{j=1}^{n_i} \varepsilon_{ij} \\ m_{i2} = n_i^{-1}\sum_{j=1}^{n_i} (\varepsilon_{ij}^2 - 1) \\ m_{i3} = n_i^{-1}\sum_{j=1}^{n_i} \left(\varepsilon_{ij}^3 - g_3(X_i)\right) \end{cases}$$

Let $m_i(\Theta) = \left[m_{i1}(\Theta) \quad m_{i2}(\Theta) \quad m_{i3}(\Theta)\right]'$, the variance covariance matrix for $m_i(\Theta)$ is

$$V(m_i(\Theta)) = \begin{bmatrix} V_{11} & V_{12} & V_{13} \\ V_{12} & V_{22} & V_{23} \\ V_{13} & V_{23} & V_{33} \end{bmatrix}$$, where

$$V_{11} = n_i^{-1}$$
$$V_{22} = n_i^{-1} V(\varepsilon_{ij}^2) = n_i^{-1}\left(E(\varepsilon_{ij}^4) - 1\right)$$
$$V_{33} = n_i^{-1} V(\varepsilon_{ij}^3) = n_i^{-1}\left(E(\varepsilon_{ij}^6) - g_3^2(X_i)\right)$$
$$V_{12} = n_i^{-1}\operatorname{cov}(\varepsilon_{ij}, \varepsilon_{ij}^2 - 1) = n_i^{-1} E(\varepsilon_{ij}^3) = n_i^{-1} g_3(X_i)$$
$$V_{13} = n_i^{-1}\operatorname{cov}\left(\varepsilon_{ij}, \varepsilon_{ij}^3 - g_3(X_i)\right) = n_i^{-1} E(\varepsilon_{ij}^4)$$
$$V_{23} = n_i^{-1}\operatorname{cov}\left(\varepsilon_{ij}^2, \varepsilon_{ij}^3 - g_3(X_i)\right) = n_i^{-1}\left(E(\varepsilon_{ij}^5) - E(\varepsilon_{ij}^3)\right) = n_i^{-1}\left(E(\varepsilon_{ij}^5) - g_3(X_i)\right)$$

The quantities $E(\varepsilon_{ij}^c)$, $c \in \{3, 4, 5\}$, can be readily derived from the moment generating function $M(t) = 2\exp(\xi t + \omega^2 t^2 / 2)\Phi(\omega\tau t)$ (Φ is the cumulative distribution function of standard normal) by taking the cth derivation and evaluating at $t = 0$, i.e.,

$E(\varepsilon_{ij}^c) = \partial M^{(c)}(t) / \partial t \,|_{t=0}$.

Appendix C: Simulation Results for 100 counties and 200 observations of each county

Coefficient	LMM (I)	M-LMM (II)	OLS (III)	GMM (IV)	IV/I	IV/II	IV/III
MC relative biases ($\times 10^{-3}$)							
Mean- α_0	-10.6082	-10.6082	-10.6082	-6.5096	0.6136	0.6136	0.6136
α_1	0.7328	0.7328	0.7328	1.4521	1.9816	1.9816	1.9816
α_2	-4.4400	-4.4400	-4.4400	-4.5129	1.0164	1.0164	1.0164
Variance- β_0	8.3610	8.3610	8.1697	-13.9225	-1.6652	-1.6652	-1.7042
β_1	1.3901	1.3901	1.4507	-7.3202	-5.2659	-5.2659	-5.0461
β_2	6.5200	6.5200	6.8661	-6.3594	-0.9754	-0.9754	-0.9262
Skewness- γ_0	1302.2075	-38.5102	-19.4708	35.1772	0.027	-0.9134	-1.8067
γ_1	-530.5507	3.1587	3.9932	8.0105	-0.0151	2.536	2.0061
γ_2	-329.5575	-6.0307	-11.1404	-23.2649	0.0706	3.8578	2.0883
MC variances ($\times 10^{-3}$)							
Mean- α_0	0.5615	0.5615	0.5615	0.4205	0.7488	0.7488	0.7488
α_1	0.0096	0.0096	0.0096	0.0080	0.8375	0.8375	0.8375
α_2	0.1645	0.1645	0.1645	0.1323	0.8042	0.8042	0.8042
Variance- β_0	4.9340	4.9340	4.9579	1.5520	0.3146	0.3146	0.3130
β_1	0.0892	0.0892	0.0897	0.0465	0.5208	0.5208	0.5179
β_2	1.2670	1.2670	1.2728	0.5169	0.4080	0.4080	0.4061
Skewness- γ_0	142.5465	3.0059	1.8747	2.1859	0.0153	0.7272	1.166
γ_1	2.4663	0.0331	0.0255	0.0316	0.0128	0.9559	1.2404
γ_2	26.1962	0.6299	0.4574	0.5386	0.0206	0.8550	1.1773
MSE ($\times 10^{-3}$)							
Mean- α_0	0.6735	0.6735	0.6735	0.4624	0.6866	0.6866	0.6866
α_1	0.0101	0.0101	0.0101	0.0101	1.0013	1.0013	1.0013
α_2	0.1842	0.1842	0.1841	0.1525	0.8287	0.8287	0.8287

Variance- β_0	4.9989	4.9989	5.0197	1.7443	0.3489	0.3489	0.3475
β_1	0.0910	0.0910	0.0917	0.1000	1.0984	1.0984	1.0903
β_2	1.3083	1.3083	1.3186	0.5569	0.4256	0.4256	0.4223
Skewness- γ_0	1838.1483	4.4859	2.2519	3.4212	0.0019	0.7627	1.5192
γ_1	283.9479	0.0430	0.0414	0.0957	0.0003	2.2260	2.3124
γ_2	134.7781	0.6657	0.5811	1.0793	0.0080	1.6214	1.8573