Testing Day's Conjecture that More Nitrogen Decreases Crop Yield Skewness

Xiaodong Du, David A. Hennessy, and Cindy L. Yu

Working Paper 10-WP 511
August 2010

Center for Agricultural and Rural Development
Iowa State University
Ames, Iowa 50011-1070
www.card.iastate.edu

Xiaodong Du is a postdoctoral research associate in the Center for Agricultural and Rural Development, David A. Hennessy is a professor in the Department of Economics and Center for Agricultural and Rural Development, and Cindy L. Yu is an assistant professor in the Department of Statistics, all at Iowa State University.

Questions or comments about the contents of this paper should be directed to Xiaodong Du, 565 Heady Hall, Iowa State University, Ames, Iowa 50011. Ph: (515) 294-8015; Fax: (515) 294-6336; E-mail: xdu@iastate.edu.

Abstract

While controversy surrounds skewness attributes of typical yield distributions, a better understanding is important for agricultural policy assessment and for crop insurance rate setting. Day (1965) conjectured that crop yield skewness declines with an increase in low levels of nitrogen use, but higher levels have no effect. In a theoretical model based on the law of the minimum (von Liebig) technology, we find conditions under which Day's conjecture applies. Employing four experimental plot datasets, we investigate the conjecture by introducing (*a*) a flexible Bayesian extension of the Just-Pope technology to incorporate skewness, and (*b*) a quantile-based measure of skewness shift. For corn yields, the Bayesian estimation provides strong evidence in favor of negative skewness at commercial nitrogen rates and for Day's conjecture. There was weaker evidence in favor of positively skewed cotton yield and little evidence in favor of the conjecture. The results are also confirmed by the quantile-based measure.

Keywords: crop insurance, Gibbs sampler, Just and Pope technology, negative skewness, quantile regression.

Risks originating from the random nature of yield have a significant impact on farmers' production and marketing decisions (Goodwin and Ker 2002). A better understanding of crop yield distributions is important both for crop producers and for the crop insurance industry, where contract payout patterns are sensitive to distribution tails. There is a long-established literature on how inputs affect mean and variance of crop yield distributions while much less is known about how inputs affect yield skewness. A positive skew indicates that the *tail* on the right side is *longer* than that on the left side and the bulk of the values lie to the left of the mean.

The best known early work on crop yield skewness was by Day (1965), whose intuition suggested positive skewness (p. 714 and p. 735). His rationale for the hypothesis of positive skewness reveals much about his views on the crop production process during the middle years of the 20th century (p. 714):

> "That field crop yields should conform to such a pattern seems plausible. Excellent weather condition (*sic*) throughout the entire growing season (particularly during the germination, flowering, heading, and harvesting season) must prevail if high yields are to be obtained. Such crop years do occur and phenomenally high yields are recorded. These pull average yields up. On the other hand, bad weather – too much or too little rain or heat – during any one of several critical periods is sufficient to reduce yields drastically, even though ideal weather is the rule during the preceding and succeeding parts of the season. Thus, common sense suggests that less than average yields are more likely than greater than average yields."

We interpret the emphasis on critical periods as a general belief that a law of the minimum production technology applies in the sense of, e.g., Paris (1992). It is noteworthy that Day explored data from Mississippi, far outside the main crop growing areas of the United States with soil and climate limitations that suggest that 'ideal' conditions are not to be expected throughout the growing season.[1] In addition, many have

[1] Soil matters in a variety of ways, including in determining how resilient a crop is to adverse weather. Deep soils high in organic matter allow deep rooting and drainage away from roots. They also ensure moisture storage across the season so that early season

argued that improved understanding of crop nutrient needs, pest control inputs (Gardner 2002) and seed genetics (Yu and Babcock 2010) have likely removed many of the factors that might constrain yield in a typical year. This is our view. We suggest further that activities intended to remove stochastic production constraints, such as increased use of nitrogen, should generally act to render skewness less positive or more negative.

Day found mixed evidence on the direction of skewness. Despite some controversy in regard to the relevance of central limit theorems and concerns about methodologies used to draw inferences from aggregated data (Just and Weninger 1999; Khoundouri and Kourogenis 2010), the preponderance of evidence since then has pointed to negative yield skewness for crops grown in the Corn Belt. See Hennessy (2009, 2010) for recent reviews. Much of the analysis has been on aggregated data, where input use is unknown.

We take Day's approach in scrutinizing input-controlled crop trial data and were inspired by some of his findings. For Mississippi cotton and corn experimental farm crop trial data at seven different nitrogen levels, he found a positive Pearson statistic for skewness. In the case of cotton, skewness tended to become less positive at higher nitrogen levels while no discernable pattern emerged for corn. Oats had positive skewness at the zero nitrogen level but negative and generally declining skewness at the higher levels. In conclusion, Day used these empirical regularities to conjecture (p. 739) that skewness decreases with increased nitrogen up to a critical level.[2]

For estimation of input effects on the crop yield distribution, one widely applied model is by Just and Pope (1978; 1979). These authors develop a stochastic production function specification that allows explicit estimation of the effect of independent

moisture can substitute in for later water deficiencies or heat-induced water stress.

[2] He does not provide any formal logical argument on why this might be so and appears to have been taken aback by them (p. 735). As we have just argued, perhaps his reasoning on why positive skewness was to be expected in mid-20th century crop production should have led him to a theoretical foundation consistent with the empirical regularities.

variables, e.g., fertilizer, on the mean and variance of the yield distribution. The 1978 paper proposes a maximum likelihood estimation (MLE) procedure while the 1979 proposes a three-step feasible generalized least square (FGLS) approach to estimate the first and second moments of stochastic yield responses. But the model does not address yield skewness.

Antle (1987) introduces a moment-based non-parametric model, which is able to express skewness as a function of inputs. But complexity and lack of efficiency limit its practical application (Yatchew 1998). More recently, Antle (2010) has proposed and implemented a partial moment regression system approach to studying the role of inputs on skewness. For real-farm potato production data in Ecuador's Northern Highland region, he finds that fertilizer use likely decreases skewness, consistent with Day's findings but in a very different context. He finds that fungicide and labor use likely increase skewness and suggests that partial moment analysis is needed to understand subtleties in how inputs affect distribution tails. Employing a two-stage MLE procedure, Nelson and Preckel (1989) propose the conditional beta distribution for crop yield. Weaknesses in their method include that (*i*) estimation efficiency is conditioned on the pre-imposed beta density, and that (*ii*) standard errors-of-moment elasticities are highly nonlinear and difficult to obtain.

The general theme of the current study is that one should expect a relaxation of production constraints to decrease yield skewness. In addition, no further effect on skewness should be expected when the input is sufficiently large that it is unlikely to constrain. We develop a theoretical framework to illustrate how more of an input should affect yield skewness and then use two approaches to investigate empirically the impact of applied nitrogen rates on skewness. In the first approach we extend the Just-Pope model to accommodate skewness. The model is then applied to several crop trial experimental datasets. Two datasets are applied to corn in Iowa, one to corn in Minnesota, and one to cotton in Texas. We conduct inference within a Bayesian framework, employing Monte Carlo Markov Chain methods. The second empirical method involves quantile regression models to estimate skewness shifts induced by nitrogen application across different portions of yield distribution.

In general, we find that corn yield skewness is positive at nitrogen levels below about 25 lb/ac, but negative at higher levels. In addition, nitrogen levels above about 75-100 lb/ac have little effect on skewness. In short, we find strong evidence in favor of Day's conjecture. For cotton, where the least observations are available, we find limited evidence of positive skewness and no evidence in favor or against Day's conjecture. After developing our theoretical model, the two empirical approaches are outlined. Then the data are explained, estimations are run and discussed, and some concluding comments are offered.

Theoretical Model

The intent of this section is to find technical conditions under which Day's conjecture applies. We will then argue that, on the whole, one should probably expect these conditions to apply. The skewness concept we adopt in this section is Pearson's standard moment-based concept, as employed in Day (1965). Start with a result due to van Zwet's (1964).[3]

Fact 1: For η random, if $h(\eta)$ is increasing and convex and the skewness statistics exist, then random variable $h(\eta)$ is more positively skewed than is η.

With $S(\eta)$ as the skewness statistic, the fact can be expressed as

[3] Theorem 2.1.1, page 10.

(1) $$S(h(\eta)) \equiv \frac{\mathbb{E}\left[\left(h(\eta)-\mathbb{E}\left[h(\eta)\right]\right)^3\right]}{\left\{\mathbb{E}\left[\left(h(\eta)-\mathbb{E}\left[h(\eta)\right]\right)^2\right]\right\}^{3/2}} \geq \frac{\mathbb{E}\left[\left(\eta-\mathbb{E}[\eta]\right)^3\right]}{\left\{\mathbb{E}\left[\left(\eta-\mathbb{E}[\eta]\right)^2\right]\right\}^{3/2}} \equiv S(\eta).$$

In Appendix A, the following is demonstrated.

Fact 2: When $h(\eta)$ is increasing and concave, then inequality (1) is reversed.

Van Zwet's observation was used by Hennessy (2010) to show a possible relationship between weather and yield distributions. Suppose that, all else equal, better weather in some index sense improves yield, but the marginal effect is decreasing. Then the weather-conditioned yield distribution will be more negatively skewed than the weather distribution itself. In the analysis to follow, our interest is in the role that market inputs play in determining skewness. More specifically, what conditions on a stochastic crop production technology would support the Day conjecture?

Following Berck and Helfand (1990), Paris (1992), Chambers and Lichtenberg (1996), and Berck, Geoghegan and Stohs (2000), the production technology is modeled as of the von-Liebig type:

(2) $$y(\xi; x+\eta) = \min[\xi, x+\eta],$$

where x is a representative input and η might be viewed as a carry-over soil endowment of that input. Variable ξ is a spatial production constraint, such as the availability of alternative nutrients in the soil, organic matter or a measure of soil compaction. Random endowment η follows distribution $F(\eta)$ with support on set $\eta \in [\underline{\eta}, \overline{\eta}]$. Endowment ξ follows mass distribution $G(\xi)$ with support $[a,b]$ across the unit of analysis, where the density exists over the entire support and where ξ is independent of η. For future reference we write the survival function as $\overline{G}(x+\eta) \equiv 1-G(x+\eta)$. Our interest is in mean, or aggregate, yield over the entire unit of analysis, i.e., integrate over space:[4]

$$
\begin{aligned}
&y(x+\eta)=\int_a^b y(\xi;x+\eta)dG(\varepsilon)=\int_a^b \min[\xi,x+\eta]dG(\varepsilon)\\
(3)\qquad &=\int_a^{x+\eta}\xi dG(\varepsilon)+(x+\eta)\bar{G}(x+\eta)=\xi G(\xi)\Big|_{\xi=a}^{\xi=x+\eta}-\int_a^{x+\eta}G(\xi)d\varepsilon+(x+\eta)\bar{G}(x+\eta)\\
&=x+\eta-\int_a^{x+\eta}G(\xi)d\varepsilon.
\end{aligned}
$$

Specify ICX and ICV as, respectively, the sets of increasing, convex and increasing, concave functions. In light of Facts 1 and 2 we wish to establish the following. For $x = x_0$ and $x = x_1$ with $x_1 > x_0$, can it be shown that either (*i*) $y(x_1+\eta)\in ICX \bigcup ICV$ when viewed as a function of $y(x_0+\eta)$, or (*ii*) $y(x_0+\eta)\in ICX \bigcup ICV$ when viewed as a function of $y(x_1+\eta)$? For (*i*) and (*ii*), it is readily shown that monotonicity applies. To see this, write $t = y(x_0+\eta)$ so that $\eta = y^{-1}(t) - x_0$ where $y^{-1}(\cdot)$ represents the inverse function of $y(\cdot)$. Then, upon substitution, we may specify $B(t)\equiv y(x_1 - x_0 + y^{-1}(t))$. The function $y(\cdot)$ is increasing so its inverse is increasing and the chain rule gives that $B(t)$ is increasing.

For convexity/concavity, the other part of each condition set in Facts 1 and 2, Cargo (1965) has shown that if $\psi(t)$ and $\chi(t)$ are both twice continuously differentiable and strictly increasing functions on domain T, then $\chi(t)$ is a convex transformation of $\psi(t)$ whenever

$$
(4)\qquad \frac{\chi_{tt}(t)}{\chi_t(t)}\ge\frac{\psi_{tt}(t)}{\psi_t(t)}\quad \forall t\in T,
$$

or, equivalently, whenever $d\,\mathrm{Ln}[\chi_t(t)]/dt \ge d\,\mathrm{Ln}[\psi_t(t)]/dt$. Under the same conditions

[4] 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.

on $\chi(t)$ and $\psi(t)$, then $\chi(t)$ is a concave transformation of $\psi(t)$ whenever the inequality in (4) is reversed. To see this, write $\chi(t) = u(\psi(t))$ and then compute $\chi_t(t) = u_\psi(\psi(t))\psi_t(t)$, $u_\psi(\psi(t)) = \chi_t(t)/\psi_t(t)$,

$u_{\psi\psi}(\psi(t))\psi_t(t) = \{\psi_t(t)\chi_{tt}(t) - \chi_t(t)\psi_{tt}(t)\}/[\psi_t(t)]^2$, and

$u_{\psi\psi}(\psi(t)) = \{\psi_t(t)\chi_{tt}(t) - \chi_t(t)\psi_{tt}(t)\}/[\psi_t(t)]^3$. The latter has the sign of $d\,\mathrm{Ln}[\chi_t(t)]/dt$ $-d\,\mathrm{Ln}[\psi_t(t)]/dt$. In our case, we need to show

(5) $$\frac{y_{\eta\eta}(x_1+\eta)}{y_\eta(x_1+\eta)} \geq \frac{y_{\eta\eta}(x_0+\eta)}{y_\eta(x_0+\eta)} \quad \forall\, \eta \in [\underline{\eta}, \overline{\eta}]$$

for an increasing and convex transformation, and also to show that the inequality is reversed for an increasing and concave transformation.

Now (3) provides $y_\eta(x+\eta) = \overline{G}(x+\eta)$ and $y_{\eta\eta}(x+\eta) = -g(x+\eta)$ so that condition (5) becomes

(6) $$\frac{g(x_1+\eta)}{\overline{G}(x_1+\eta)} \leq \frac{g(x_0+\eta)}{\overline{G}(x_0+\eta)} \quad \forall\, \eta \in [\underline{\eta}, \overline{\eta}].$$

Given that $x_1 > x_0$, this is a monotonicity condition. It requires that $\overline{G}(x+\eta)g_\eta(x+\eta) +$ $[g(x+\eta)]^2 \leq 0$. An alternative way of stating it is to write $J^G(x+\eta) \equiv \mathrm{Ln}[\overline{G}(x+\eta)]$ and then

(7) $$J^G_{xx}(x+\eta) \geq 0 \quad \forall\, \eta \in [\underline{\eta}, \overline{\eta}].$$

The log of the survival function needs to be convex; this attribute is often referred to as the log-convex survival function property. Bagnoli and Bergstrom (2005) in their Theorem 4 and Table 3 have this to be true for the Weibull, Pareto and Gamma distributions under certain shape parameter conditions. In that case, an increase in nitrogen use would make skewness more positive.

A third way of characterizing (6) is that hazard ratio $g(x+\eta)/\bar{G}(x+\eta)$ is decreasing. To explain a decreasing hazard rate in our context, consider the ratio's numerator and denominator. Numerator $g(x+\eta)$ can be viewed as the probability that an additional unit of input x is marginal, i.e., will have no effect on production. Denominator $\bar{G}(x+\eta)$ conditions the distribution to provide the probability that factor $x+\eta$ is marginal in the sense that this factor has limited yield. Thus a decreasing hazard ratio increases the probability that an additional unit of the input is effective in increasing production given that lower levels have been effective.[5] This seems to be rather unreasonable.[6]

Condition (5) is reversed, i.e., an increasing and concave transformation applies, whenever

(8) $$J_{xx}^{G}(x+\eta) \le 0 \quad \forall\, \eta \in [\underline{\eta}, \overline{\eta}],$$

or the survival function is log-concave. This is true whenever $\bar{G}(x+\eta)g_{\eta}(x+\eta) + [g(x+\eta)]^2 \ge 0$. The condition applies whenever the density function is concave. More generally, it applies whenever the density function is log-concave, or $g(\cdot)g_{\eta\eta}(\cdot) \le [g_{\eta}(\cdot)]^2$ (An 1998). As such, it applies for the normal, uniform, logistic, and the extreme value distribution with distribution function $\exp[-\exp(-\xi)]$ among others of interest to economists (An 1998).[7] We consider this more likely than a log-convex survival function

---

5 Given the productivity interpretation just provided, a comment on (5) may prove helpful. It is apparent from (5) that the matter at issue is relative curvature where the best-known application in economics is attributable to Pratt (1964). There, the topic was the effect on degree of risk aversion as wealth moves along a utility function. Here, the topic is the effect on input expected productivity as the input level moves along the survival function that depicts how effective the input is in expectation.

6 The analogy in the actuarial science or reliability statistics literatures would be that the probability of living to birthday 86 given that one has lived to birthday 85 is larger than the probability of living to birthday 85 conditional on having lived to birthday 84.

7 The statistics literature on log-concave densities, distributions and survival functions is very large. We refer the reader to Dharmadhikari and Joag-dev (1988) for an extensive

as it requires the conditional probability the input is effective at the margin to be declining in the amount of input.

In conclusion, we can state Richard Day's conjecture as follows:

Proposition 1: If the survival function for the mass distribution of spatial endowment factor ξ is log-concave, or (8) applies, then an increase in input x makes average production over the unit of analysis less positively or more negatively skewed. If the survival function is log-convex, or (7) applies, then the reverse is true.

Of course, if market input x has price $w > 0$ and this price increases, then the law of factor demand would have an increase in input use (p. 131 in Chambers 1988). Under (8), the result would be a more negative skew on the yield distribution.

Corollary 1: Suppose (*i*) the survival function for spatial endowment factor ξ is log-concave and (*ii*) the standard law of factor demand applies. Then, *ceteris paribus*, a decrease in the price of the market input will lead to a less positive or more negative skew on average production over the unit of analysis. If, instead, the survival function is log-convex, then the reverse is true.

A comment is in order in regard to simulations in table 1 of Hennessy (2009) where a pair of random variables (ω_1, ω_2) are jointly normally distributed and output is $y =$ $\min[\omega_1, \omega_2]$. There, an increase in the mean of one might increase or decrease the skewness of the output distribution. The unit of analysis there is at one spatial point. Integrating out over one of the random variables, when considering aggregate yield

---

but dated review.

where one variable is viewed as having a mass density over space, smoothes the skewness statistic so that a uniformly monotone effect can be identified under certain regularity conditions. These are conditions (7) and (8).

From an economic perspective, our general reading of the proposition and corollary is as follows. The real price of crop nutrients has in the main decreased dramatically over the past century (Gardner 2002; Federico 2005). All else fixed, our result identifies conditions under which this should affect the nature of skewness in a definite way. In what follows we will test for how inputs have affected skewness.

Empirical Analysis

We will use two distinct approaches to provide evidence on how inputs affect yield skewness. The first extends the Just-Pope specification to account for skewness and uses Bayesian methods to implement the approach. The second method invokes a quantile regression to study how the quantile gaps stretch or contract as the input changes.

Bayesian analysis of a skewness measure

Let experimental plot crop yield be given by

(9) $$y = f(z) + g(z)\varepsilon^{h(z)}$$

where z is an input, e.g., nitrogen. Here ε is random and $f(\cdot)$, $g(\cdot)$ and $h(\cdot)$ are functions to be parameterized and estimated. To be consistent with Just-Pope method applications, let

(10) $$f(z) = \alpha_0 z^{\alpha_1}; \qquad g(z) = \beta_0 z^{\beta_1}; \qquad h(z) = \gamma_0 z^{\gamma_1}.$$

The moments for crop yield are

$$\mathbb{E}[y] = \alpha_0 z^{\alpha_1} + \beta_0 x^{\beta_1} \mathbb{E}[\varepsilon^{\gamma_0 z^{\gamma_1}}]; \quad \mathbb{E}\left[(y - \mathbb{E}[y])^2\right] = \beta_0^2 z^{2\beta_1} \mathrm{Var}\left(\varepsilon^{\gamma_0 z^{\gamma_1}}\right);$$

$$(11) \qquad S(y) = \frac{\mathbb{E}\left[\left(\varepsilon^{\gamma_0 z^{\gamma_1}} - \mathbb{E}[\varepsilon^{\gamma_0 z^{\gamma_1}}]\right)^3\right]}{\left\{\mathbb{E}\left[\left(\varepsilon^{\gamma_0 z^{\gamma_1}} - \mathbb{E}[\varepsilon^{\gamma_0 z^{\gamma_1}}]\right)^2\right]\right\}^{3/2}}.$$

Eqn. (11) implies that (*i*) skewness is determined by (γ_0, γ_1) only, (*ii*) scale (and so yield variance) is determined by $(\beta_0, \beta_1, \gamma_0, \gamma_1)$ so that scale can change with the level of z independent of skewness, and (*iii*) location (and so yield mean) is determined by $(\alpha_0, \alpha_1, \beta_0, \beta_1, \gamma_0, \gamma_1)$ so that location can change with z independent of yield variance and skewness. In summary, the stochastic production function in Eqn. (9) is mean, variance and skewness flexible.

We estimate the stochastic production technology using crop nitrogen trial data. Given the discrete and limited number of nitrogen levels applied in such trials, we estimate the impacts of the nitrogen input on the skewness of crop yield distribution for each nitrogen level individually. For the nitrogen application level i, $i \in \Omega_I \equiv \{1, 2..., I\}$, we adopt Eqn. (9) to give our empirical crop yield model as

$$(12) \qquad y^i = a_0^i + \mathbf{X}\boldsymbol{\beta} + b^i \varepsilon^{1/c^i}, \;\; \varepsilon \sim D \times Beta(\alpha, \alpha); \;\; i \in \Omega_I.$$

where $y^i = (y_1^i, y_2^i, \ldots, y_k^i)'$ denotes the *k* plot-level crop yield observations for the *i*th nitrogen level. Parameter a_0^i denotes the constant term in the yield equation where it is allowed to vary with nitrogen application, and it includes the effect of nitrogen on mean yield, $f(z^i)$. Matrix $\mathbf{X} = (x_1 \; x_2 \; \ldots \; x_L)$, in which $x_l = (x_{l1}, x_{l2}, \ldots, x_{lI})' \;\; \forall l \in \Omega_L$, denotes controlled variables such as location, rotation, and/or tillage effect, and technological innovation of crop production is represented by a time dummy or trend. The corresponding coefficient vector is $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_L)'$. The scale effect in specification

(9) is represented by $b^i = g(z^i)$. The probability density function (pdf) of the general beta distribution is $f(x;\alpha_1,\alpha_2) = \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} x^{\alpha_1-1}(1-x)^{\alpha_2-1} = \frac{1}{B(\alpha_1,\alpha_2)} x^{\alpha_1-1}(1-x)^{\alpha_2-1}$, where $\Gamma(\cdot)$ is the gamma function and the beta function $B(\alpha_1,\alpha_2)$ appears as a normalization constant. The above pdf is defined on the interval $(0,1)$ with two positive shape parameters, α_1 and α_2. Instead we employ a symmetric beta distribution, $Beta(\alpha,\alpha)$, which is parameterized by $\alpha(=\alpha_1=\alpha_2)$.

We specify the yield random variations as ε^{1/c^i} with skewness parameter c^i, and assume that ε follows a symmetric beta distribution $Beta(\alpha,\alpha)$ on the range $[0,D]$. Based on the constructive representation in (12), skewness is introduced into an originally symmetric distribution on ε through parameter c^i. Doing so allows us to retain some well-known properties of symmetric distributions, so that Fact 1 and Fact 2 allow skewness to be ordered in a common framework. Notice that the transformation $f(\varepsilon) = \varepsilon^{1/c^i}$ is increasing and convex. The symmetric specification of ε in Eqn. (12), together with Fact 1 and Fact 2, means that condition $c^i < 1$ (i.e., $1/c^i > 1$) implies a positive yield skew at the *i*th nitrogen level. On the other hand, condition $c^i > 1$ (i.e., $1/c^i < 1$) implies a negative skew. In addition, higher c^i is associated with decreasing skewness.

Inference here is conducted within a Bayesian framework. One important advantage of adopting the Bayesian approach is that it is relatively easy to incorporate inequality constraints on parameters into the estimation procedure. The inequality constraint is to ensure that parameter estimates are consistent with relationships implied by the underlying distribution assumption, which is $\frac{1}{D}\left[(y^i - a_0^i - \mathbf{X\beta})/b^i\right]^{c^i} \in (0,1)$, i.e.,

$0 < \frac{1}{D}\left[(y^i - a_0^i - \mathbf{X\beta}) / b^i \right]^{c^i} < 1$. Furthermore, the Bayesian estimation procedure is particularly suitable as the model specified in Eqn. (12) is very nonlinear in the parameters, which makes it very difficult for the MLE estimator to converge.

For the parameter vector in model (12), $\Theta = \{\mathbf{a_0}, \boldsymbol{\beta}, \mathbf{b}, \mathbf{c}, D, \alpha\}$, where $\mathbf{a_0} = \{a_0^i\}_{i=1}^I$, $\mathbf{b} = \{b^i\}_{i=1}^I$, $\mathbf{c} = \{c^i\}_{i=1}^I$, Bayesian inference stems from the joint posterior distribution of the model parameters conditional on the observed data, which can be expressed as follows:

$$p(\mathbf{a_0}, \boldsymbol{\beta}, \mathbf{b}, \mathbf{c}, D, \alpha \mid Y, \mathbf{X})$$

$$(13) \quad \propto \left[\frac{1}{D^\alpha B(\alpha,\alpha)} \right]^N \prod_{i=1}^{I} \prod_{k=1}^{n} \frac{c^i}{b^i} \left(\frac{y_k^i - a^i}{b^i} \right)^{c^i \alpha - 1} \left[1 - \frac{1}{D} \left(\frac{y_k^i - a^i}{b^i} \right)^{c^i} \right]^{\alpha - 1} \times I_A \times p(\Theta);$$

$$a^i = a_0^i + \mathbf{X\beta}; \quad I_A = \begin{cases} 1 \ \text{if} \ \frac{1}{D}\left(\frac{y_k^i - a^i}{b^i} \right)^{c^i} \in (0,1), \\ 0 \qquad \text{otherwise}; \end{cases}$$

where $N (= i \times n)$ are the total observation numbers and $B(\alpha,\alpha)$ $(= \int_0^1 t^{\alpha-1}(1-t)^{\alpha-1} dt)$ denotes the beta function. The first three items are the likelihood function, while the last term in the product, $p(\Theta)$, denotes the joint prior distribution of model parameters.

The Gibbs sampler, one of the Monte Carlo Markov Chain methods in Bayesian implementation, is applied by repeated sampling from the conditional posterior density of each parameter. It is obvious that the posterior distribution for each parameter derived from the joint posterior distribution in Eqn. (13) doesn’t belong to any convenient distribution family. Thus we employ the random-walk Metropolis-Hasting algorithms (see the detailed introduction and application in, e.g., Gelman et al. 2007, and Koop, Poirier and Tobias 2007) for updating draws from each posterior distribution of model parameters. The technical details regarding the conditional posterior distributions and implementation of the Gibbs sampler are provided in Appendix B. After convergence, the draws of the Gibbs sampler for each parameter are used to compute the mean and standard deviation of the marginal posterior distribution.

Quantile regression analysis of skewness shifts

Quantile-based measures can be used to describe shape-shifts of a distribution including skewness (Hao and Naiman 2007, Ch. 2). Note that for a skewed distribution, the quantile function $Q^{(p)}$, which is defined as $P(Z \le Q^{(p_i)}) = p_i$ for a random variable Z and probability $p_i \in [0,1]$, is asymmetric around the median. Based on a random sample, the quantile-based measure, $S^{(p)}$, defined as the ratio of the spreads above and below the median (upper spread vs. lower spread), $S^{(p)} = [(Q^{(1-p)} - Q^{(0.5)}) / (Q^{(0.5)} - Q^{(p)}) - 1]$ for $p < 0.5$ could be used to measure skewness (Hao and Naiman 2007).[8] For symmetric distributions, $Q^{(1-p)} - Q^{(0.5)} = Q^{(0.5)} - Q^{(p)}$ so that $S^{(p)} < 0$ means the concave relation in *p*, $Q^{(1-p)} - Q^{(0.5)} < Q^{(0.5)} - Q^{(p)}$, which is consistent in spirit with Fact 2.

Compared with a reference case, disproportional upper and lower spread changes relative to the median indicate skewness shifts in a comparison case. Thus the sample-based skewness shift, $SS^{(p)}$, for the middle $100(1-2p)\%$ of the population is defined as

$$(14) \qquad SS^{(p)} = \left[\frac{(Q_C^{(1-p)} - Q_C^{(0.5)}) / (Q_R^{(1-p)} - Q_R^{(0.5)})}{(Q_C^{(0.5)} - Q_C^{(p)}) / (Q_R^{(0.5)} - Q_R^{(p)})} - 1 \right] \text{ for } p < 0.5$$

(Hao and Naiman 2007),[9] where $Q_{C(R)}^{(p)}$ is p quantile for the comparison (reference) case.

In this study, the quantile regression model (QRM) introduced by Koenker and Bassett (1978) is employed to characterize the impact of explanatory variables, e.g.,

---

[8] Eqn. 2.2, page 14.

[9] Eqn. 5.2, page 72. We don't calculate sample-based skewness shift $SS^{(p)}$ in this study. Instead, we employ and report model-based measure $SKS^{(p)}$ at a later juncture.

nitrogen, on the shape of the yield distribution. The model is specified as

(15) $$y^i = X'\gamma_p = \alpha + \gamma_z z^i + X'_{\{-z\}}\gamma_{\{-z\},p}, \quad \text{Quan}_p(y^i \mid X) = X'\gamma_p, \quad i \in \Omega_I$$

where $\text{Quan}_\theta(y^i \mid X) = X'\gamma_\theta$ denotes the conditional quantile of crop yield, y^i, on a set of controlled variables, $\mathbf{X} = (z\ X_{\{-z\}})$, in which the first variable is the nitrogen application rates, z, while $X'_{\{-z\}} = (x_2\ \ldots\ x_L)$ are other controlled variables as defined in Eqn. (12). All controlled variables are centered at sample means.

Following the measure in Eqn. (14), the skewness shift resulting from additional nitrogen application can be calculated from the estimated QRM coefficients[10]

(16) $$SKS^{(p)} = 1000\left[\frac{(\hat{\gamma}_z^{(1-p)} + \hat{\alpha}^{(1-p)} - \hat{\gamma}_z^{(0.5)} - \hat{\alpha}^{(0.5)}) / (\hat{\alpha}^{(1-p)} - \hat{\alpha}^{(0.5)})}{(\hat{\gamma}_z^{(0.5)} + \hat{\alpha}^{(0.5)} - \hat{\gamma}_z^{(p)} - \hat{\alpha}^{(p)}) / (\hat{\alpha}^{(0.5)} - \hat{\alpha}^{(p)})} - 1\right] \text{ for } p < 0.5,$$

where the estimated constant $\hat{\alpha}$'s provide the fitted quantile functions at the means of independent variables and are defined as the reference cases. Standard errors of the estimated skewness shifts can be obtained by the bootstrap method (Greene 2003, p. 924). For a given portion, $100(1-2p)\%$, of yield population, $SKS^{(p)} < 0$ indicates a reduction of right-skewness induced by an increase of nitrogen application. Conversely, $SKS^{(p)} > 0$ indicates an exacerbation of right-skewness, i.e., the crop yield is more positively/right skewed.

Data

In this study, four crop yield datasets—A, B, C, and D—are employed. We do not consider Day's data because, apart from the different methodologies we bring to bear, that would amount to in-sample validation of the conjecture. One common feature of the datasets is that all have been collected from controlled experiments, although the

[10] Eqn. 5.2, page 72. We have introduced the factor 1000, as quantile bound $p \in [0,1]$ leads to very small numbers.

experiments are conducted under different control conditions such as tillage practices and rotation sequences.

Dataset A: Corn production data are obtained from controlled experiments conducted at Iowa State University's Research and Demonstration Farm located in Floyd County, Iowa, from 1979 to 2003 (Mallarino, Ortiz-Torres and Pecinovsky 2004). The corn yield data are collected under four rotations, <C>, <CS>, <CCS>, and <CCCS>, where <CCCS> is to be read as the corn-corn-corn-soybeans rotation. Four levels of nitrogen treatments, 0 lb/ac, 80 lb/ac, 160 lb/ac, and 240 lb/ac are applied for corn. Each combination of rotation and nitrogen level is replicated three times each a year. Using the same dataset, Hennessy (2006) investigates the rotation effect on crop yield and on nitrogen input choices. Rotation effect is found to persist for one year. Taking the one-year memory into account, dataset A contains 525 observations for each nitrogen level. The corresponding specification for Eqn. (12) after including the controlled variables is

$$y^i = a_0^i + \sum_{j=1}^{2} x_j \beta_j + b^i \varepsilon^{1/c^i}, \;\; \varepsilon \sim D \times Beta(\alpha, \alpha); \;\; i = 1, 2, 3, 4$$

where $\boldsymbol{\beta} = [\beta_1, \beta_2]'$ represents the average effects of crop rotation <CC> and linear time trend on corn yields.

Dataset B: Corn yields in dataset B are generated by controlled experiments on fifteen geographically dispersed Iowa farms under <CC> and <CS> rotations from 1985 to 1990 under ten levels of nitrogen fertilizer, 0 lb/ac, 25 lb/ac, 50 lb/ac, 75 lb/ac, 100 lb/ac, 120 lb/ac, 150 lb/ac, 200 lb/ac, 250 lb/ac, and 300 lb/ac. In the context of yield and revenue insurance, Babcock and Hennessy (1996) employ part of this dataset to estimate conditional distributions of crop yields. There are three replications of each nitrogen level. Total observations for each level of nitrogen vary from 193 to 203. The control variables for Eqn. (12) include five dummies, $\beta_1,...,\beta_5$, for the years 1985, 1986, 1987, 1988, and 1989, one dummy β_6 for the CC rotation, and fourteen dummy variables for farm locations ($\beta_7,...,\beta_{14}$). The corresponding specification is

$$y^i = a_0^i + \sum_{j=1}^{20} x_j\beta_j + b^i\varepsilon^{1/c^i},\ \ \varepsilon \sim D\times Beta(\alpha,\alpha);\ \ i=1,2,...,10.$$

Dataset C: In dataset C, corn yield data are collected from replicated plot experiments in Morris County, Minnesota, from 1984 to 1989 with five nitrogen fertilizer rates of 0 lb/ac, 40 lb/ac, 80 lb/ac, 120 lb/ac, and 160 lb/ac.[11] For each nitrogen rate, there are 167 or 168 observations. Besides the time indicators for 1985-1989 ($\beta_1,...,\beta_4$), dummies for various tillage practices, fall plow (FP; β_5), fall chisel (FC; β_6), and ridge till (RD; β_7), are also included as controlled variables in the Eqn. (12) regression, which is specified as

$$y^i = a_0^i + \sum_{j=1}^{7} x_j\beta_j + b^i\varepsilon^{1/c^i},\ \ \varepsilon \sim D\times Beta(\alpha,\alpha);\ \ i=1,2,...,5.$$

Dataset D: We choose to consider cotton yield data because the crop has received some attention in the skewness literature. Ramirez, Mishra and Field (2003) studied dryland cotton in the West Texas Plains counties of Childress, Cochran, Crosby, Hale and Wichita. Like Day, these authors identify positive skewness. Similar to Day's intuition, they ascribe it to the positive skewness of the limiting factor, rainfall. Our cotton yield data are obtained from field experiments in Calhoun, San Patricio, and Wharton counties in the Texas Coastal Bend cotton growing region for 1998-2002.[12] Nitrogen fertilizer is

---

[11] See Mitchell (2004) for more details on the dataset. We thank Paul Mitchell for providing us with datasets C and D.

[12] The original data do not provide irrigation information. Comparing the yields in the data to the NASS county yields and typical dryland vs. irrigated yields in the literature, it

applied at four levels, 0 lb/ac, 50 lb/ac, 100 lb/ac, and 150 lb/ac. The controlled variables in Eqn. (12) are two location dummies for the first two sampling areas, β_1 and β_2, and time dummies for the years 1998-2001 ($\beta_3,...,\beta_6$). There are 39 yield observations for each nitrogen rate. The regression model is specified as

$$y^i = a_0^i + \sum_{j=1}^{6} x_j\beta_j + b^i\varepsilon^{1/c^i}, \ \varepsilon \sim D\times Beta(\alpha,\alpha); \ \ i=1,2,...,4.$$

Parameter Estimates and Discussion of Results

Bayesian analysis of a skewness measure

The Gibbs sampler is fully documented in Appendix B and is coded in Matlab. We run the posterior simulator for 10,000 simulations and discard the first 5,000 simulations as burn-in. Estimation results on generated data experiments revealed that the Bayesian algorithm is able to recover parameters of the data generation process well. The priors of the model parameters are set to be reasonably non-informative as follows: $a_0^i \sim N(10,10)$; $b^i \sim N(100,20)$; $c^i \sim N(1,1)$; $D \sim N(0,10)$; $\alpha \sim N(0,10)$; and $\beta_i \sim N(0,10)$. For each of datasets A, B, C, and D, Appendix C presents coefficient posterior means, standard deviations and model parameter probabilities of being positive.

Following Eqn. (11), the skewness of conditional corn yield distribution for each nitrogen application level is calculated as

is clear that the yields are dryland cotton. We thank Paul Mitchell for pointing this out. See Seo, Mitchell and Leatham (2005) for more details on the dataset.

$$S(y^i) = \frac{\mathbb{E}\left[\left(\varepsilon^{d^i} - \mathbb{E}[\varepsilon^{d^i}]\right)^3\right]}{\left\{\mathbb{E}\left[\left(\varepsilon^{d^i} - \mathbb{E}[\varepsilon^{d^i}]\right)^2\right]\right\}^{3/2}}$$

(17)

$$= \frac{\frac{B(\alpha+3d^i,\alpha)}{B(\alpha,\alpha)} - \frac{3B(\alpha+2d^i,\alpha)B(\alpha+d^i,\alpha)}{[B(\alpha,\alpha)]^2} + 2\left(\frac{B(\alpha+d^i,\alpha)}{B(\alpha,\alpha)}\right)^3}{\frac{B(\alpha+2d^i,\alpha)}{B(\alpha,\alpha)} - \left(\frac{B(\alpha+d^i,\alpha)}{B(\alpha,\alpha)}\right)^2},$$

where $d^i = 1/c^i$. The mean and standard errors of skewness for each nitrogen level are calculated using the draws from posterior distributions of individual parameters and are presented in table 1.

It can be seen from the results in table 1 that for datasets A, B, and C the estimated skewness measures are generally consistent with our theoretical results and prior expectation. An increase in nitrogen inputs makes average corn production more negatively skewed. While positive skewness is shown at zero or low levels of nitrogen rates, higher levels of nitrogen reduce skewness to negative values. Skewness ceases to decrease further in response to an increase in nitrogen at levels more than approximately 100 lb/ac.

Specifically, in dataset A, we see strong evidence of negative skewness associated with non-zero nitrogen rates. The posterior standard deviations for the skewness parameters are quite small relative to their means. An increase in nitrogen is associated with an increase in the absolute values of skewness. In datasets B and C, we see similar patterns for skewness shifts, but the inference from these data are not as strong given that the posterior standard deviations are of nearly equal magnitude to the posterior means.

For cotton yield in dataset D, positive skewness is found for all nitrogen rates. The finding is consistent with findings in the literature (e.g., Day 1965; Ramirez, Misra and

Field 2003). As pointed out in some other studies, the positive skewness of cotton yield could be related to rainfall and drought conditions in the sampling areas. In other words, nitrogen is generally not the limiting factor.

Quantile regression analysis of skewness shifts

Estimation results of the quantile regression model are reported at nine quantiles ranging from 0.025 to 0.975 in Appendix D. All explanatory variables are centered at sample means. Based on the estimated coefficients, skewness shifts in Eqn. (16) are calculated and their standard errors are computed from 1,000 bootstrap samples. Table 2 presents the means and standard errors of the estimated skewness shifts for the middle $100(1-2p)\%$ of the conditional crop yield distribution.

The results in table 2 indicate that higher levels of nitrogen application slightly decrease right-skewness for all selected measures in datasets A and B. Estimated $SKS^{(p)}$ values in dataset A are significantly different from zero at the 1% significance level, while in dataset B the negative effects are only highly significant for the middle 95% and 97.5% of the population. The percentage decrease in skewness induced by one more unit of nitrogen range from 0.17% to 0.22% in dataset A and from 0.02% to 0.21% in dataset B. On average, the magnitude of impact appears to be smaller in dataset B.

Although not significant in general, the impacts in dataset C are not consistent across different portions of the yield population, which may be because of the relatively smaller sample size. For the middle 50% and 90% of the yield population, the effect of high nitrogen application is negative, i.e., reducing right-skewness, but the effect changes to positive for 95% and 97.5% of the population. A positive and small effect of nitrogen on cotton yield skewness is found in dataset D, indicating that a more right-skewed yield distribution is associated with higher levels of nitrogen.

Conclusion

Working with the law of the minimum technology, we provide a modeling framework in which to interpret Day's (1965) observations that more nitrogen tends to make skewness less positive or more negative, but only up to a point. This framework allows us to find reasonable conditions on the distribution of a competing constraint (e.g., soil characteristics) such that aggregate skewness decreases with more of a (possibly) constraining input. We develop two approaches to assess the role of nitrogen in determining yield skewness. One is a generalization of the Just-Pope technology, implemented with Bayesian methods, while the other is a quantile regression approach.

Experimental plot datasets allow us to address the typical concerns of temporal and spatial aggregation in yield modeling. For corn yields, estimation results from both methods provide strong evidence that negative skewness is associated with non-zero nitrogen rates. In addition, more negative skewness is associated with more nitrogen, and some evidence is provided that a ceiling nitrogen level exists above which skewness does not change. There is weaker evidence for positively skewed cotton yield, and we found no discernable skewness pattern as nitrogen levels change.

In conclusion, we think that development of a fuller theory on how input constraints affect yield skewness will have to await empirical regularities emerging from studies on diverse crops grown in different production environments. Although perhaps most convenient to work with, nitrogen is not the only management practice that can be varied in a controlled manner. For example, controlled experimental data on conservation tillage practices are available (e.g., DeVuyst and Halvorson 2004). To the extent that conservation tillage promotes soil water storage, one might expect an effect on yield skewness in water-constrained cropping areas. Data on irrigation effects are also available (e.g., Eck 1984; Kim et al. 2008), where one would expect that an increase in availability of irrigation water will act to make yield skewness less positive or more negative.

References

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

Antle, J.M. 1987. "Econometric Estimation of Producers' Risk Attitudes." *American Journal of Agricultural Economics* 69(3):509–522.

Antle, J.M. 2010. "Asymmetry, Patrial Moments and Production Risk." *American Journal of Agricultural Economics* (forthcoming).

Babcock, B.A., and D.A. Hennessy. 1996. "Input Demand under Yield and Revenue Insurance." *American Journal of Agricultural Economics* 78(2):416–427.

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

Berck, P., J. Geoghegan, and S. Stohs. 2000. "A Strong Test of the von Liebig Hypothesis." *American Journal of Agricultural Economics* 82(4):948–955.

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

Cargo, G.T. 1965. "Comparable Means and Generalized Convexity." *Journal of Mathematical Analysis and Applications* 12(3):387–392.

Chambers, R.G. 1988. *Applied Production Analysis: A Dual Approach*. Cambridge UK: Cambridge University Press.

Chambers, R.G., and E. Lichtenberg. 1996. "A Nonparametric Approach to the von Liebig-Paris Technology." *American Journal of Agricultural Economics* 78(2):373–386.

Day, R.H. 1965. "Probability Distributions of Field Crop Yields." *Journal of Farm Economics* 47(3):713–741.

DeVuyst, E.A., and A.D. Halvorson. 2004. "Economics of Annual Cropping versus Crop-Fallow in the Northern Great Plains as Influenced by Tillage and Nitrogen." *Agronomy Journal* 96(1):148–153.

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

Eck, H.V. 1984. "Irrigated Corn Yield Response to Nitrogen and Water." *Agronomy Journal* 76(3):421–428.

Federico, G. 2005. *Feeding the World: An Economic History of Agriculture, 1800-2000*. Princeton NJ: Princeton University Press.

Gardner, B.L. 2002. *American Agriculture in the Twentieth Century: How It Flourished and What It Cost*. Cambridge MA: Harvard University Press.

Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin. 2004. *Bayesian Data Analysis*, 2nd edition. London: Chapman & Hall/CRC.

Goodwin, B.K., and A.P. Ker. 2002. "Modeling Price and Yield Risk." In R.E. Just and R.D. Pope ed. *A Comprehensive Assessment of the Role of Risk in U.S. Agriculture*. Norwell, MA: Kluwer Academic Publisher, Ch. 14.

Greene, W.H. 2003. *Econometric Analysis*. Upper Saddle River, NJ: Prentice Hall.

Hao, L., and D.Q. Naiman. 2007. *Quantile Regression*. Thousand Oaks, CA: SAGE Publications.

Hennessy, D.A. 2006. "On Monoculture and the Structure of Crop Rotations." *American Journal of Agricultural Economics*, 88(4):900–914.

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

Hennessy, D.A. 2010. "Crop Yield Skewness and the Normal Distribution." *Journal of Agricultural and Resource Economics* 33(1):34–52.

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

Just, R.E., and R.D. Pope. 1979. "Production Function Estimation and Related Risk Considerations." *American Journal of Agricultural Economics* 61(2):277–284.

Just, R.E., and Q. Weninger. 1999. "Are Crop Yields Normally Distributed?" *American Journal of Agricultural Economics* 81(2):287–304.

Khoundouri, P., and N. Kourogenis. 2010. "On the Distribution of Crop Yields: Does the Central Limit Theorem Apply?" DEOS Working Paper #1007, Athens University of Economics and Business.

Kim, K.-I., D.E. Clay, C.G. Carlson, S.A. Clay, and T. Trooien. 2008. "Do Synergistic Relationships between Nitrogen and Water Influence the Ability of Corn to Use Nitrogen Derived from Fertilizer and Soil?" *Agronomy Journal* 100(2):551–556.

Koenker, R., and G. Bassett. 1978. "Regression Quantiles." *Econometrica* 46(1):33–50.

Koop, G., D.J. Poirier, and J.L. Tobias. 2007. *Bayesian Econometric Methods*. Cambridge, UK: Cambridge University Press.

Mallarino, A.P., E. Ortiz-Torres, and K. Pecinovsky. 2004. "Effects of Crop Rotation and Nitrogen Fertilization on Crop Production." Annual Progress Report, Northeast Research and Demonstration Farm. Iowa State University.

Mitchell, P. 2004. "Nutrient Best Management Practice Insurance and Farmer Perceptions of Adoption Risk." *Journal of Agricultural and Applied Economics* 36(3): 657–673.

Nelson, C.H., and P.V. Preckel. 1989. "The Conditional Beta Distribution as a Stochastic Production Function." *American Journal of Agricultural Economics* 71(2):370–378.

Paris, Q. 1992. "The von Liebig Hypothesis." *American Journal of Agricultural Economics* 74(4):1019–1028.

Pratt, J.W. 1964. "Risk Aversion in the Small and in the Large." *Econometrica* 32(1-2):122–136.

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

Seo, S., P. Mitchell, and D.J. Leatham. 2005. "Effects of Federal Risk Management Programs on Optimal Acreage Allocation and Nitrogen Use in a Texas Cotton-Sorghum System." *Journal of Agricultural and Applied Economics* 37(3): 685-699.

Yatchew, A. 1998. "Nonparametric Regression Techniques in Economics." *Journal of Economic Literature* 36(2):669–721.

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

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

Table 1. Estimated mean and standard deviation (in parentheses) for derived skewness

Dataset A					
Nitrogen	0	80	160	240	
Skewness	0.1317 (0.0685)	-0.1458 (0.0453)	-0.1715 (0.0544)	-0.1739 (0.0478)	
Dataset B					
Nitrogen	0	25	50	75	100
Skewness	0.1571 (0.0821)	0.0562 (0.0618)	-0.0559 (0.0677)	-0.1537 (0.0893)	-0.1793 (0.0842)
Nitrogen	125	150	200	250	300
Skewness	-0.1931 (0.0870)	-0.1626 (0.0913)	-0.1958 (0.0886)	-0.1719 (0.887)	-0.1972 (0.0884)
Dataset C					
Nitrogen	0	40	80	120	160
Skewness	0.2073 (0.0868)	-0.0909 (0.0752)	-0.0728 (0.0867)	-0.1316 (0.0807)	-0.0079 (0.0975)
Dataset D					
Nitrogen	0	50	100	150	
Skewness	0.1403 (0.0850)	0.0290 (0.0721)	0.0735 (0.0664)	0.0998 (0.0717)	

Table 2. Estimated skewness shifts for the middle $100(1-2p)\%$ of conditional yield distribution

Dataset A				
p	0.25	0.1	0.05	0.025
$S^{(p)}$	0.00	-0.16	-0.17	-0.17
$SKS^{(p)}$	-2.2***	-2.0***	-1.7***	-2.0***
	(0.7)	(0.5)	(0.4)	(0.5)
Dataset B				
p	0.25	0.1	0.05	0.025
$S^{(p)}$	-0.14	-0.34	-0.41	-0.42
$SKS^{(p)}$	0.2	-1.2**	-1.7***	-2.1***
	(0.8)	(0.6)	(0.5)	(0.5)
Dataset C				
p	0.25	0.1	0.05	0.025
$S^{(p)}$	0.19	0.34	0.33	0.36
$SKS^{(p)}$	-0.1	-0.7	1.1	1.0
	(2.0)	(1.7)	(1.5)	(1.4)
Dataset D				
p	0.25	0.1	0.05	0.025
$S^{(p)}$	-0.07	0.12	0.37	0.40
$SKS^{(p)}$	1.6	1.7	1.6	1.5
	(3.9)	(2.6)	(2.5)	(2.4)

Note: Single (*), double (**), and triple (***) asterisks denote significance at 0.10, 0.05, and 0.01 levels, respectively.

Appendix A. *Demonstration of Fact 2*

Define $g(x) = -h(-x)$. It is readily shown that $g(x)$ is increasing and concave in x if and only if $h(\cdot)$ is increasing and concave in its argument. For any arbitrary random variable ω, define $\zeta \equiv -\omega$ so that ζ is also a random variable. Inequality (1) holds for any random variable, so we may write $S(g(\zeta)) \geq S(\zeta)$. Hence, by substitution,

$$\begin{aligned}
&S(-h(-\zeta)) \geq S(\zeta) \\
&\Rightarrow -S(h(-\zeta)) \geq -S(-\zeta) \\
&\Rightarrow S(h(-\zeta)) \leq S(-\zeta) \\
&\Rightarrow S(h(\omega)) \leq S(\omega).
\end{aligned} \tag{A1}$$

The first implication follows from the fact that $\mathbb{E}[(\eta - \mathbb{E}[\eta])^3] = -\mathbb{E}[(-1)^3(\eta - \mathbb{E}[\eta])^3] = -\mathbb{E}[(-\eta - \mathbb{E}[-\eta])^3]$. The final inequality follows from direct substitution in which how ω impacts $h(\cdot)$ is not relevant. ■

Appendix B. *Conditional Posterior Distribution and Gibbs Sampler*

As specified in Eqn. (13), parameters a_0^i, b^i, and c^i are associated with the i th nitrogen level only, so they are estimated from the corresponding set of yield observations, y_k^i, $k \in \{1, \ldots, n_i\}$. While the yield observations at all nitrogen levels contain information on the coefficients $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_L)'$, and the distribution parameters D and α, in the following Gibbs sampling algorithm, they are estimated based on the whole sample set with observation $n = \sum_{i=1}^{I} n_i$.

Step 1. Drawing from the conditional posterior pdf of a_0^i:

$$p(a_0^i \mid \cdot) \propto \prod_{k=1}^{n_i} \left(\frac{y_k^i - a^i}{b^i} \right)^{c^i \alpha - 1} \left[1 - \frac{1}{D} \left(\frac{y_k^i - a^i}{b^i} \right)^{c^i} \right]^{\alpha - 1} \times p(a_0^i) \times I_A, \; i \in \Omega_I.$$

Here the last term $p(a_0^i)$ denotes the prior information on parameter a_0^i and I_A is the inequality constraint defined in Eqn. (13). Imposition of the inequality constraint I_A is implemented within the random-walk Metropolis-Hastings algorithm (and similarly in the following steps). So in each loop, the potential new parameter value generated from the random walk process needs to satisfy both the typical updating rule and the inequality constraint to be accepted as a new update in the parameter space.

Step 2. Drawing from the conditional posterior pdf of b^i:

$$p(b^i \mid \cdot) \propto \left(\frac{1}{b^i} \right)^{n_i} \prod_{k=1}^{n_i} \left(\frac{y_k^i - a^i}{b^i} \right)^{c^i \alpha - 1} \left[1 - \frac{1}{D} \left(\frac{y_k^i - a^i}{b^i} \right)^{c^i} \right]^{\alpha - 1} \times p(b^i) \times I_A, \; i \in \Omega_I.$$

Step 3. Drawing from the conditional posterior pdf of c^i:

$$p(c^i \mid \cdot) \propto \left(c^i \right)^{n_i} \prod_{k=1}^{n_i} \left(\frac{y_k^i - a^i}{b^i} \right)^{c^i \alpha - 1} \left[1 - \frac{1}{D} \left(\frac{y_k^i - a^i}{b^i} \right)^{c^i} \right]^{\alpha - 1} \times p(c^i) \times I_A, \; i \in \Omega_I.$$

Step 4. Drawing from the conditional posterior pdf of β_l :

$$p(\beta_l \mid \cdot) \propto \prod_{i=1}^{I}\prod_{k=1}^{n}\left(\frac{y_k^i - a^i}{b^i}\right)^{c^i\alpha-1}\left[1-\frac{1}{D}\left(\frac{y_k^i - a^i}{b^i}\right)^{c^i}\right]^{\alpha-1} \times p(\beta_l)\times I_A,\ l \in \Omega_L.$$

Step 5. Drawing from the conditional posterior pdf of D :

$$p(D \mid \cdot) \propto \left(\frac{1}{D}\right)^{\alpha N}\prod_{i=1}^{I}\prod_{k=1}^{n}\left(\frac{y_k^i - a^i}{b^i}\right)^{c^i\alpha-1}\left[1-\frac{1}{D}\left(\frac{y_k^i - a^i}{b^i}\right)^{c^i}\right]^{\alpha-1} \times p(D)\times I_A.$$

Step 6. Drawing from the conditional posterior pdf of α :

$$p(\alpha \mid \cdot) \propto \left(\frac{1}{D^{\alpha}B(\alpha)}\right)^{N}\prod_{k=1}^{n}\left(\frac{y_k^i - a^i}{b^i}\right)^{c^i\alpha-1}\left[1-\frac{1}{D}\left(\frac{y_k^i - a^i}{b^i}\right)^{c^i}\right]^{\alpha-1} \times p(\alpha)\times I_A.$$

Appendix C. *Bayesian Estimation Results on Datasets A, B, C, and D*

Table C1. Estimation results on dataset A

Variable	$E(\cdot \mid y)$	$Std(\cdot \mid y)$	$\Pr(\cdot > 0 \mid y)$	Variable	$E(\cdot \mid y)$	$Std(\cdot \mid y)$	$\Pr(\cdot > 0 \mid y)$
a_0^1	11.63	5.35	1.00	c^3	1.30	0.11	1.00
b^1	52.51	8.21	1.00	a_0^4	13.71	6.12	1.00
c^1	0.85	0.07	1.00	b^4	114.43	12.05	1.00
a_0^2	11.43	6.33	1.00	c^4	1.31	0.09	1.00
b^2	96.43	10.47	1.00	D	2.78	0.30	1.00
c^2	1.25	0.08	1.00	α	5.94	0.85	1.00
a_0^3	15.78	6.62	1.00	β_1	-29.33	1.54	0.00
b^3	108.68	8.50	1.00	β_2	0.16	0.11	0.93

β_1 : dummy for rotation effect of <CC>; β_2 : time variable 1-25 for 1979-2004.

Table C2. Estimation results on dataset B

Variable	$E(\cdot \mid y)$	$Std(\cdot \mid y)$	$\Pr(\cdot > 0 \mid y)$	Variable	$E(\cdot \mid y)$	$Std(\cdot \mid y)$	$\Pr(\cdot > 0 \mid y)$
a_0^1	19.69	8.55	1.00	c^9	1.34	0.19	1.00
b^1	62.97	16.59	1.00	a_0^{10}	14.97	7.43	1.00
c^1	0.81	0.10	1.00	b^{10}	116.41	15.15	1.00
a_0^2	23.98	7.93	1.00	c^{10}	1.41	0.20	1.00
b^2	72.42	15.25	1.00	D	3.15	1.30	1.00
c^2	0.92	0.08	1.00	α	8.08	2.71	1.00
a_0^3	24.78	9.04	1.00	β_1	-10.82	10.13	0.03
b^3	85.57	14.60	1.00	β_2	6.58	11.10	0.90
c^3	1.10	0.12	1.00	β_3	-2.09	13.06	0.52
a_0^4	26.57	9.03	1.00	β_4	-58.71	14.73	0.00
b^4	95.72	18.67	1.00	β_5	-8.97	12.67	0.02
c^4	1.30	0.18	1.00	β_6	-27.95	12.13	0.00
a_0^5	17.92	9.64	1.00	β_7	-2.90	8.73	0.43
b^5	107.01	18.85	1.00	β_8	15.18	13.77	0.91
c^5	1.36	0.18	1.00	β_9	-33.53	7.52	0.00
a_0^6	16.40	7.08	1.00	β_{10}	14.77	10.00	0.94
b^6	112.86	14.51	1.00	β_{11}	-1.81	10.75	0.49
c^6	1.40	0.19	1.00	β_{12}	5.23	8.85	0.78
a_0^7	22.32	7.26	1.00	β_{13}	-1.20	12.68	0.56
b^7	106.94	14.57	1.00	β_{14}	20.45	15.57	0.93
c^7	1.32	0.19	1.00	β_{15}	7.93	13.04	0.80
a_0^8	16.35	7.24	1.00	β_{16}	-1.94	13.48	0.54
b^8	116.38	14.21	1.00	β_{17}	-30.02	13.18	0.00
c^8	1.41	0.20	1.00	β_{18}	-9.14	17.77	0.37
a_0^9	16.35	7.63	1.00	β_{19}	-3.14	7.94	0.36
b^9	115.19	18.19	1.00	β_{20}	-3.14	15.65	0.95

$\beta_1 - \beta_5$: dummies for 1985-1989; β_6: dummy for <CC> rotation. $\beta_7 - \beta_{20}$: location dummies.

Table C3. Estimation results on dataset C

Variable	$E(\cdot \mid y)$	$Std(\cdot \mid y)$	$\Pr(\cdot > 0 \mid y)$	Variable	$E(\cdot \mid y)$	$Std(\cdot \mid y)$	$\Pr(\cdot > 0 \mid y)$
a_0^1	3.64	2.81	1.00	a_0^5	32.63	11.72	1.00
b^1	58.93	10.45	1.00	b^5	84.80	14.57	1.00
c^1	0.76	0.09	1.00	c^5	1.03	0.15	1.00
a_0^2	10.40	5.93	1.00	D	2.43	0.28	1.00
b^2	87.18	10.93	1.00	α	8.24	1.69	1.00
c^2	1.17	0.14	1.00	β_1	-47.75	2.95	0.00
a_0^3	25.77	8.59	1.00	β_2	-35.91	2.60	0.00
b^3	86.63	11.92	1.00	β_3	3.13	2.36	0.90
c^3	1.14	0.16	1.00	β_4	-70.19	3.21	0.00
a_0^4	18.48	8.30	1.00	β_5	5.43	2.49	0.98
b^4	95.56	12.61	1.00	β_6	4.69	2.43	0.97
c^4	1.26	0.17	1.00	β_7	1.89	2.29	0.79

$\beta_1 - \beta_4$: dummies for 1985, 1986, 1987, and 1988.

$\beta_5 - \beta_7$: dummies for tillage practices FP, FC, and RD.

Table C4. Estimation results on dataset D

Variable	$E(\cdot \mid y)$	$Std(\cdot \mid y)$	$\Pr(\cdot > 0 \mid y)$	Variable	$E(\cdot \mid y)$	$Std(\cdot \mid y)$	$\Pr(\cdot > 0 \mid y)$
a_0^1	254.54	46.89	1.00	b^4	48.79	10.31	1.00
b^1	37.54	10.84	1.00	c^4	0.88	0.08	1.00
c^1	0.83	0.10	1.00	D	26.46	6.17	1.00
a_0^2	324.17	52.78	1.00	α	7.09	2.32	1.00
b^2	54.53	9.83	1.00	β_1	-513.37	191.22	0.00
c^2	0.96	0.10	1.00	β_2	-303.86	226.08	0.04
a_0^3	320.09	45.01	1.00	β_3	-225.86	222.23	0.19
b^3	51.84	11.19	1.00	β_4	-18.55	161.96	0.56
c^3	0.91	0.08	1.00	β_5	0.93	78.77	0.54
a_0^4	289.35	41.65	1.00	β_6	2.13	70.22	0.53

$\beta_1 - \beta_2$: dummies for location a and b; $\beta_3 - \beta_6$: dummies for 1998-2001.

Appendix D. *Quantile Regression Results on Datasets A, B, C and D*

Table D1. Quantile regression results on dataset A

Dataset A	Quantile								
Variable	0.025	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.975
β_z	0.23^{w}	0.243^{w}	0.234^{w}	0.28^{w}	0.31^{w}	0.30^{w}	0.31^{w}	0.30^{w}	0.27^{w}
β_1	-24.69^{w}	-23.10^{w}	-27.58^{w}	-31.90^{w}	-31.21^{w}	-30.61^{w}	-28.59^{w}	-24.19^{w}	-24.40^{w}
β_2	-0.31^{u}	-0.16	-0.19^{u}	-0.10	0.26^{v}	0.33^{w}	0.55^{w}	0.48^{w}	0.60^{w}
α	61.52^{w}	68.51^{w}	77.17^{w}	95.77^{w}	120.83^{w}	142.60^{w}	158.64^{w}	167.29^{w}	174.83^{w}

β_1 : dummy for rotation effect of <CC>; β_2 : time variable 1-25 for 1979-2004.

Note: *u*, *v* and *w* denote significance at 0.10, 0.05, and 0.01 levels, respectively.

Table D2. Quantile regression results on dataset B

Dataset B	Quantile								
Variable	0.025	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.975
β_z	0.11[v]	0.13[w]	0.14[w]	0.16[w]	0.15[w]	0.15[w]	0.11[w]	0.10[w]	0.08[w]
β_1	15.41	16.87[u]	9.17	-3.82	-22.75[w]	-19.52[w]	-13.42[v]	-2.40	-20.34[v]
β_2	-14.49	-2.94	-5.90	-2.79	4.40	13.19[w]	14.81[w]	16.54[w]	16.81[w]
β_3	2.92	5.24	-3.10	-3.70	-2.79	-2.81	-2.43	1.21	-2.12
β_4	-68.58[w]	-68.25[w]	-75.21[w]	-74.14[w]	-57.03[w]	-48.83[w]	-36.67[w]	-29.97[w]	-34.96[w]
β_5	-14.18[w]	-10.89[w]	-18.35[w]	-19.66[w]	-6.93[w]	-4.68[u]	-1.91	-1.56	-1.84
β_6	-29.03[w]	-27.99[w]	-24.23[w]	-24.50[w]	-26.84[w]	-25.38[w]	-21.22[w]	-16.38[w]	-18.53[w]
β_7	15.52	4.54	12.86	22.13[w]	19.02[w]	26.73[w]	31.46[w]	23.79[w]	26.14[w]
β_8	-22.74[u]	16.33[w]	29.23[w]	40.24[w]	42.66[w]	49.83[w]	55.66[w]	51.52[w]	49.84[w]
β_9	3.98	-13.30	-7.05[w]	-1.73	-18.36[w]	-27.65[w]	-39.26[w]	-52.33[w]	-46.58[w]
β_{10}	20.77	10.35	24.84[w]	40.66[w]	50.04[w]	47.14[w]	45.20[w]	36.64[w]	51.03[w]
β_{11}	-4.27	-15.25	-5.90	7.24	30.76[w]	33.81[w]	33.23[w]	27.17[w]	26.49[v]
β_{12}	16.48	6.79	25.56[w]	37.71[w]	35.81[w]	27.99[w]	24.99[w]	17.31[w]	19.21[u]
β_{13}	-2.83	-6.54	7.68	20.48[w]	27.55[w]	31.23[w]	31.77[w]	26.47[w]	27.12[w]
β_{14}	25.76[u]	14.80[u]	27.36[w]	38.01[w]	44.07[w]	52.74[w]	60.38[w]	61.80[w]	61.39[w]
β_{15}	-5.08	-12.50	8.85	29.58[w]	39.36[w]	44.45[w]	50.07[w]	52.17[w]	55.25[w]
β_{16}	10.05	-1.79	11.38[u]	25.06[w]	29.56[w]	24.61[w]	38.94[w]	40.14[w]	41.61[w]
β_{17}	-20.32	-34.18[w]	-22.35[w]	-7.44	-5.10	5.41	11.31[w]	13.33[w]	13.75
β_{18}	28.75[v]	10.74	16.09	21.95[w]	17.00[v]	26.87[w]	16.63[v]	12.59[u]	11.23
β_{19}	16.97	22.98[u]	29.65[w]	28.76[w]	17.44[w]	14.69[v]	13.73[u]	8.10	14.90[u]

β_{20}	51.02^{w}	45.82^{w}	52.59^{w}	54.68^{w}	57.62^{w}	57.54^{w}	60.87^{w}	54.53^{w}	51.43^{w}
α	91.72^{w}	89.85^{w}	102.19^{w}	125.14^{w}	142.36^{w}	163.47^{w}	176.92^{w}	176.19^{w}	183.14^{w}

$\beta_1 - \beta_5$: dummies for 1985-1989; β_6 : dummy for <CC> rotation. $\beta_7 - \beta_{20}$: location dummies.

Table D3. Quantile regression results on datasets C and D

Dataset C	Quantile								
Variable	0.025	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.975
β_z	0.31[v]	0.33[w]	0.33[w]	0.31[w]	0.27[w]	0.25[w]	0.26[w]	0.27[w]	0.28[w]
β_1	-33.20[u]	-31.50[w]	-35.60[w]	-40.75[w]	-50.71[w]	-62.70[w]	-66.60[w]	-68.90[w]	-69.45[w]
β_2	-31.70[u]	-27.40[w]	-33.40[w]	-34.90[w]	-40.50[w]	-46.80[w]	-39.10[w]	-44.73[w]	-43.90[w]
β_3	-3.20	1.60	3.60	3.45	5.24[u]	-5.00	-3.10	-2.70	0.50
β_4	-64.50[w]	-63.50[w]	-67.00[w]	-70.90[w]	-72.89[w]	-85.40[w]	-87.80[w]	-91.13[w]	-84.20[w]
β_5	9.80	3.15	6.10u	5.50[v]	0.73	3.10	4.80	10.53[u]	12.30[v]
β_6	4.20	4.60	7.50[v]	11.95[w]	4.20[u]	2.50	1.10	4.83	6.30
β_7	3.20	2.00	1.50	4.45	1.36	0.80	2.10	8.50	7.20
α	45.29[w]	50.62[w]	57.09[w]	72.58[w]	89.91[w]	104.77[w]	118.44[w]	126.48[w]	132.03[w]

$\beta_1 - \beta_4$: dummies for 1985, 1986, 1987, and 1988. $\beta_5 - \beta_7$: dummies for tillage practices FP, FC, and RD.

Dataset D	Quantile								
Variable	0.025	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.975
β_z	0.75[u]	0.70	0.62	0.90 [w]	0.58v	0.82v	0.87	1.04	0.76[w]
β_1	75.33	86.80	119.57[v]	95.50[u]	121.00[w]	202.50[w]	264.33[w]	278.00[w]	250.00[w]
β_2	102.00[w]	110.60[u]	129.29[w]	162.75[w]	211.00[w]	332.83[w]	381.33[w]	395.00[w]	337.00[w]
β_3	-379.00[w]	-349.00[w]	-303.57[w]	-394.00[w]	-441.00[w]	-526.17[w]	-671.33[w]	-663.00[w]	-677.00[w]
β_4	-215.00[w]	-206.40[w]	-198.57[w]	-224.50[w]	-169.00[w]	-201.83[w]	-283.67[v]	-243.00[w]	-49.00[w]
β_5	21.67	18.80	7.71	-63.50	-116.00	-243.50[w]	-430.00[w]	-445.00[w]	-429.00[w]
β_6	-17.00	-17.00	-21.00	-109.25[u]	-120.00[u]	218.17[w]	272.33[u]	389.00[w]	409.00[w]
α	610.07[w]	621.14[w]	645.06 [w]	738.16 [w]	830.86 [w]	979.87[w]	1074.01[w]	1126.39[w]	1166.87[w]

$\beta_1 - \beta_2$: dummies for location a and b; $\beta_3 - \beta_6$: dummies for 1998-2001.