Crop Yield Skewness under the Law of Minimum Technology

David A. Hennessy

Working Paper 07-WP 451
July 2007

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

David Hennessy is a professor in the Department of Economics and Center for Agricultural and Rural Development at Iowa State University.

Questions or comments about the contents of this paper should be directed to David Hennessy, 578C Heady Hall, Iowa State University, Ames, IA 50011-1070; Ph.: (515) 294-6171; Fax: (515) 294-6336; E-mail: hennessy@iastate.edu.

Abstract

A large empirical literature exists seeking to identify crop yield distributions. Consensus has not yet formed. This is in part because of data aggregation problems but also in part because no satisfactory motivation has been forwarded in favor of any distribution, including the normal. This article explores the foundations of crop yield distributions for the Law of the Minimum, or weakest-link, resource constraint technology. It is shown that heterogeneity in resource availabilities can increase expected yield. The role of stochastic dependence is studied for the technology. With independent, identical, uniform resource availability distributions the yield skew is positive, whereas it is negative whenever the distributions are normal. Simulations show how asymmetries in resource availabilities determine skewness. Extreme value theory is used to suggest a negative yield skew whenever production is in a tightly controlled environment so that the left tails of resource availability distributions are thin.

Keywords: beta-normal distribution, crop insurance, extreme value theory, Liebig technology, limiting factors, order statistics, reliability, weakest link.

JEL classification: Q1, D2, D8

Introduction

Two major and unresolved themes in the production economics of crop agriculture concern responses to inputs absent uncertainty and yield distributions conditional on inputs. Nature, through sunshine, rainfall, and other weather variables, ensures that crop inputs are stochastic. In addition, inputs applied during cultivation do not equate with inputs available to the plant, and this is in part because of weather-dependent linkages involving soil temperature, soil biological activity, and run-off. Thus, these two themes cannot be separated in that if one does not understand input-output relations absent uncertainty about input availabilities then one cannot know much about these relations in the presence of uncertainty. The intent of this article is to seek firmer footing regarding the structural foundations of yield distributions. In doing so, we will pay particular attention to one controversial feature of yield distributions, namely, crop yield skewness.

To further these goals, a stance must be taken on the deterministic structure of crop production technologies. Although long controversial, the only technology with clearly motivated foundations is the Sprengel and von Liebig "law of the minimum," henceforth referred to as LoM. The idea is that crop input availabilities are perfect complements such that the most limiting resource determines output; e.g., $y = \min[a^1(x_1), a^2(x_2), \ldots]$, where resource availability $a^i(x_i)$ is a non-decreasing function of some input x_i. When the $a^i(x_i)$ functions are linear then the technology is referred to as a linear LoM. In general, the technology requires that surplus resource availabilities (RAs) have a null marginal product.[1] When advocating the technology,

---

[11] The weakest-link technology also arises in the economics of financing public goods (Hirshleifer 1983; Cornes 1993), as well as in health economics (Dow, Philipson, and Sala-I-Martin 1999).

von Liebig famously suggested the analogy with what is now referred to as the Liebig barrel. This is a barrel with a regular bottom but where staves have different lengths at the top. Capacity is determined by the shortest stave so that lengthening any other stave has no effect. The form is a generalization of Leontief's fixed-proportions technology specification. The claim has found some, but limited, empirical support.

Paris (1992) used a widely studied Iowa corn production experiment data set to find support for a non-linear LoM specification, where Frank, Beattie, and Embleton (1990) had earlier found evidence against a linear version with that data set. Using a dual approach and nonparametric data envelopment methods on (again) Iowa corn data, Chambers and Lichtenberg (1996) find mixed results on input substitutability consequences of the specification. Llewelyn and Featherstone (1997) used a simulation approach to identify evidence in favor of a non-linear LoM specification. Berck, Geoghegan, and Stohs (2000) took a nonparametric regression approach to test for the absence of input substitution to find little support for LoM.

In the agronomy literature, Cerrato and Blackmer (1990) are among a large number who have favored the specification. Others, as in Bloom, Chapin, and Mooney (1985), Chapin *et al.* (1987), Rastetter and Shaver (1992), Sinclair and Park (1993), and Lynch and Ho (2005), point to a multiple limitation hypothesis, or MLH. This hypothesis uses an economic framework and views nutrients as currency to be allocated within the plant to suggest that "growth is equally limited by all resources" (Bloom, Chapin, and Mooney 1985, p. 367). Taking an evolutionary economics perspective, the general tenet of this rapidly expanding literature is that successful plant species (i.e., survivors) are the genetic variants that best support biological pathways to substitute for limiting resources when at risk. For example, one means of effecting substitution is to store nutrients (at a cost) for possible later use. Laboratory tests, as in Rubio, Zhu, and Lynch

(2003), of these alternative hypotheses on a wide range of nutrients are not conclusive. It appears that the LoM is appropriate for many nutrient comparisons while for others the production process is more involved.

Thus, research is quite inconclusive on the LoM. A potential reason for this, at least for some data sets considered above, is the role of spatial non-uniformities in the production setting. Berck and Helfand (1990) have pointed out that integration over such non-uniformities can smooth over non-differentiable points in an LoM technology so that the observed noisy data may rationalize an alternative response technology. Our modeling framework will assume a generalized LoM technology in the presence of noise when seeking to understand crop yield distributions.

The literature on yield distributions, though not as extensive, is also unresolved. As with identifying the nature of a deterministic technology, the complexity of a biological system requires careful conditioning of the environment to test for technical attributes. Even under experimental conditions, field cropping is far from ideal in this regard. Parallel to the LoM, there also exists a yield distribution that is advocated by reference to theoretical foundations, namely, the normal. Here, the idea in the background is often that yield realizations over a sufficiently large area will differ because of many distinct shocks. So, the reasoning goes, some central limit theorem can be invoked to identify the normal as the limiting distribution.

The most widely cited early work on yield distributions is that of Day (1965). His data were from nitrogen-conditioned experimental cotton, corn, and oat plots in Mississippi over the middle part of the twentieth century. While finding strong evidence in favor of positive skewness (i.e., loosely where the bulk of the probability mass is to the left of the mean) for cotton, there was weaker evidence in favor of positive skewness for corn and fairly strong evidence in favor of negative skewness for oats. In addition, his skewness estimates tended to decline with an increase in the nitrogen application rate for each of the three crops.[2] This suggests that RA constraints are important in determining skewness.

A sample of more recent studies includes Gallagher (1987), for U.S.-level soybean yields over 1941-84, who found evidence of negative skewness. Nelson and Preckel (1989) and Nelson (1990), for farm-level commercial corn in Iowa over 1961-70, suggested negative skewness, as did work by Swinton and King (1991) on Minnesota commercial corn production over 1944-87. Moss and Shonkwiler (1993) have found negative skewness when analyzing U.S.-level corn yield data over 1930-90. However, Just and Weninger (1999) have emphasized methodology concerns with much of this large body of work. Data aggregation across space and possible misspecification of control factors (including time) may occur. In addition, they have expressed concerns about how significance tests on normality, the typical reference distribution, had been interpreted and/or presented for interpretation.

Endeavoring to control for these criticisms, Ramirez, Mishra, and Field (2003) have identified negative skewness for Iowa corn and soybeans using annual average data over 1950-99, and positive skewness for Texas Plains dryland cotton data, 1970-99. Sherrick *et al.* (2004), for University of Illinois data 1992-99, have subsequently found very suggestive evidence for negative skewness in corn and soybean yields. In conclusion, although the methodologies may have been remiss in certain ways, the variety in crop data sets studied, years of observation, and methods used suggest the existence of non-zero skewness. For midwestern corn and soybeans and for more recent data, the preponderance of evidence points strongly to negative skewness.

---

[2] See his Table 3 on p. 722. His work was also noteworthy in suggesting the use of the beta distribution as one sufficiently flexible in moment range to model input-conditioned distribution functions.

This article will address the technical implications of the LoM technology in the presence of stochastic RAs. It will be shown that there is reason to believe that the inputs affecting RAs will be economic complements whether or not the RAs are statistically independent. This means that an increase in the crop's price will increase all input choices and an increase in the price of any input will decrease all input choices. The implications for producer profit of different stochastic dependence structures are also explored to find that stronger positive dependence between RAs should increase expected profit for any given vector of input choices. Notwithstanding what the weakest-link technology might suggest about the technical cost of heterogeneous RAs, we identify cases where expected yield should increase with heterogeneity in availabilities, all else equal.

Turning to skewness, three statistical models of RAs are considered, where in each case the distributions of availabilities are controlled to have null skew in order to avoid introducing bias. The distributions considered are the uniform, the normal, and the raised cosine. It is shown that positive or negative skewness in yield can be supported. Analysis and simulation methods are used to explore how heterogeneity in the means and variances of RAs act to modify yield skewness. Heterogeneity in means tend to marginalize the contribution of some RAs so that the statistical attributes of the others, including skewness, determine yield distribution attributes. Contraction in the variance of one RA can also affect yield skewness in a well-defined manner. It can mass probability toward the upper end of a yield distribution and so may promote negative skewness. An increase in correlation among RAs tends to reduce the relevance of the LOM constraints because the likelihood increases that just one RA dominates as a constraint on production.

It is argued too that the motive for the empirical observation of typically negative skew in crop yields for prime agricultural cropland may be, in part, sourced in a limiting distribution law.

But a central limit theorem for the first-order (i.e., least-order) statistic may be a more appropriate reference point than the standard central limit theorems for means. If RA distributions have thin left tails, then the LoM suggests, together with extreme value theory, a bias toward negative skewness. In intensively cultivated areas where most inputs can be controlled with some precision, one might expect thin left tails on the RA distributions and thus negatively skewed yield.

Framework

The LoM yield technology for RAs $\varepsilon_i \in [0, \varepsilon_i^u] \subset \mathbb{R}_+$, $i \in \Omega_N \equiv \{1, 2, \dots, N\}$, asserts a yield realization as

(1) $\quad y = \min[\varepsilon_1, \dots, \varepsilon_N].$

This expression is very general because the distribution of $\varepsilon = (\varepsilon_1, \dots, \varepsilon_N) \subset \mathbb{R}_+^N$ is determined by the market input vector $x \in \overline{\mathbb{R}}_+^M$, among other factors, where $\overline{\mathbb{R}}_+^M$ is the positive closed M-dimensional orthant of reals. Market inputs are enumerated as x_j, $j \in \{1, 2, \dots, M\} \equiv \Omega_M$.

The upper bound on each RA, ε_i^u, is assumed fixed for convenience as it will not be relevant to our analysis, and we define $y^u = \max_{i \in \Omega_N} \{\varepsilon_i^u\}$. The unit output price is P. Factor prices are w_j where $w \in \overline{\mathbb{R}}_+^M$ represents the vector of factor prices.

If the ε_i are random, then (1) provides the first-order, or least-order, statistic (David and Nagaraja 2003). Model the ε_i as *independent* with input-conditioned distributions $F^i(\varepsilon_i \mid x)$, continuously differentiable in ε_i and twice continuously differentiable in x. Survival functions are $\overline{F}^i(\varepsilon_i \mid x) = 1 - F^i(\varepsilon_i \mid x)$, and the general formula for the cumulative distribution of y is

$$G(y\,|\,x) = \text{Prob}[\varepsilon_1 \le y \text{ or } \ \dots\ \text{ or } \varepsilon_N \le y\,|\,x] = 1 - \text{Prob}[\varepsilon_1 > y, \varepsilon_2 > y, \dots, \varepsilon_N > y\,|\,x]$$
$$= 1 - \bar{G}(y\,|\,x); \qquad \bar{G}(y\,|\,x) = \prod_{j=1}^{N} \bar{F}^j(y\,|\,x). \tag{2}$$

Its probability density function is

$$g(y) = \bar{G}(y\,|\,x) \sum_{i=1}^{N} \frac{f^i(y\,|\,x)}{\bar{F}^i(y\,|\,x)}, \tag{3}$$

while expected profit is

$$V(x;P,w) = PE[y\,|\,x] - \sum_{k=1}^{M} w_k x_k = P\int_0^{y^u} \bar{G}(y\,|\,x)dy - \sum_{k=1}^{M} w_k x_k, \tag{4}$$

where we have computed the input-conditioned expected yield as

$$E[y\,|\,x] = \int_0^{y^u} y dG(y\,|\,x) = yG(y\,|\,x)\Big|_0^{y^u} - \int_0^{y^u} G(y\,|\,x)dy = \int_0^{y^u} \bar{G}(y\,|\,x)dy. \tag{5}$$

Thus, the optimality conditions are

$$P\int_0^{y^u} \bar{G}(y\,|\,x) \sum_{i=1}^{N} \left(\frac{\partial \bar{F}^i(y\,|\,x)/\partial x_k}{\bar{F}^i(y\,|\,x)} \right) dy - w_k = 0 \quad \forall k \in \Omega_M, \tag{6}$$

with solution arguments $x_k^*(P,w)$.

Our first point is that complementarity is preserved under weak conditions. The cross-derivatives with respect to x_k and x_l are

$$\begin{aligned} \frac{\partial V(x;P,w)}{\partial x_k \partial x_l} &= P\int_0^{y^u} \bar{G}(y\,|\,x) \sum_{i=1}^{N} \left(\frac{\partial^2 \bar{F}^i(y\,|\,x)/\partial x_k \partial x_l}{\bar{F}^i(y\,|\,x)} \right) dy \\ &+ P\int_0^{y^u} \bar{G}(y\,|\,x) \sum_{i=1}^{N} \left(\frac{\partial \bar{F}^i(y\,|\,x)/\partial x_k}{\bar{F}^i(y\,|\,x)} \right) \sum_{\substack{r=1,\\ r\neq i}}^{N} \left(\frac{\partial \bar{F}^r(y\,|\,x)/\partial x_l}{\bar{F}^r(y\,|\,x)} \right) dy. \end{aligned} \tag{7}$$

This is positive $\forall k,l \in \Omega_M, k \neq l$, so long as (*i*) the inputs induce (weakly) a first-order dominating shift in an RA, or $\partial F^i(y\,|\,x)/\partial x_k \le 0\ \forall y \in \bar{\mathbb{R}}_+, \forall i \in \Omega_N, \forall k \in \Omega_M$, and (*ii*) each RA distribution is (again weakly) submodular in inputs, or $\partial^2 F^i(y\,|\,x)/\partial x_k \partial x_l \le 0\ \forall y \in \bar{\mathbb{R}}_+$,

$\forall i \in \Omega_N, \forall k, l \in \Omega_M, k \neq l$. Thus, $V(x;P,w)$ is supermodular in the vector of market inputs since any twice continuously differentiable function is supermodular whenever all second-order cross derivatives are non-negative. For a supermodular function with constant unit input costs and a constant unit output price, Theorem 10 (p. 166) in Milgrom and Shannon (1994) shows that the inputs complement in the economic sense and inputs are normal in the output price.

LEMMA 1: *For a LoM technology where the input-conditioned RAs are independent, let (i) an increase in any input induce (weakly) a first-order dominating shift in each marginal RA distribution, and (ii) these marginal distributions be (weakly) submodular in inputs. Then all inputs decrease with either an increase in any input price or a decrease in the output price.*

In particular, condition (*ii*) certainly applies when each input is dedicated to a single resource availability (e.g., irrigated water for the water resource and artificial fertilizer for the nitrogen resource) because then $\partial^2 F^i(y\,|\,x)/\partial x_k \partial x_l \equiv 0 \;\forall y \in \overline{\mathbb{R}}_+, \forall i \in \Omega_N, \forall k, l \in \Omega_M, k \neq l$. In general, non-positive cross-derivatives with respect to inputs on the independent cumulative marginals ensure supermodularity on expected output because ε_i is an increasing function of itself and $\partial^2\{\int_0^{\varepsilon_i^u} \varepsilon_i dF^i(\varepsilon_i\,|\,x)\}/\partial x_j \partial x_k = \partial^2\{\int_0^{\varepsilon_i^u} \bar{F}^i(\varepsilon_i\,|\,x) d\varepsilon_i\}/\partial x_j \partial x_k = -\int_0^{\varepsilon_i^u} \{\partial^2 F^i(\varepsilon_i\,|\,x)/\partial x_j \partial x_k\} d\varepsilon_i$ under very general real analysis conditions.

Lemma 1 begs the following question. If the LoM applies and there is RA uncertainty, then must the market inputs complement? That is, can the independence assumption be relaxed? The answer is in the affirmative.

PROPOSITION 1: *Assume the crop yield survival function* $\bar{G}(y\,|\,x)$ *is twice continuously* *differentiable in* $x \in \overline{\mathbb{R}}_+^M$ *with* $F^i(y \mid x)$ *as the marginal distributions for RAs. Assume (i) and (ii) in Lemma 1, but make no assumptions on the dependence structure between marginals. If the LoM applies, then inputs must be economic complements.*

The proof is provided in the appendix. Thus, under mild smoothness requirements, the complementarity attribute of the deterministic LoM technology is shown to be robust to the introduction of uncertainty and even arbitrary structure on how the marginals interact. Clearly, the first-order dominance requirement cannot be relaxed. The differentiability assumptions could be relaxed with little consequence, but the analysis would become cumbersome without the convenience of differential operations.

A definition allows us to make a further point with (4), one concerning the technology alone.

DEFINITION 1: (Lehmann 1966, p. 1137; Shaked and Shanthikumar 2007, p. 392) *For* $\omega \in \mathbb{R}^N$ *and* $I \subseteq \mathbb{R}^N$*, suppose a cumulative distribution* $J(\omega): I \to [0,1]$ *has marginals* $J^i(\omega_i)\ \forall i \in \Omega_N$*. If* $J(\omega) \geq \prod_{i=1}^N J^i(\omega_i)\ \forall \omega \in I$ *and* $\bar{J}(\omega) \geq \prod_{i=1}^N \bar{J}^i(\omega_i)\ \forall \omega \in I$*, then the distribution is said to be positive quadrant dependent, or PQD.*

When compared with independence, and considering only two dimensions, the definition requires a larger probability mass to the southwest of any given point, and also a larger probability mass to the northeast of that point, too. The stochastic ordering is intended to measure the extent of covariability between the set of random variables, and one implication is that $\mathrm{Cov}(\omega_i, \omega_j) \geq 0\ \forall i, j \in \Omega_N$.

PROPOSITION 2: *Suppose the technology is LoM and the marginal distributions for input-*

conditioned RAs are fixed. If the joint input-conditioned RA distribution is PQD along the hyper-line $\varepsilon_i = \lambda \; \forall i \in \Omega_N$ *then expected output and expected producer profit is larger than were the input-conditioned RAs independent.*

To confirm this, set $y = \lambda = \varepsilon_i \; \forall i \in \Omega_N$, substitute yield survival function $\prod_{j=1}^{N} \bar{F}^j(y \mid x)$ into (4) and compare with $\bar{F}(y \mid x)$ at each y realization where $F(y \mid x)$ is PQD. The proposition asserts that PQD among the RAs increases expected output, when compared with independence. For any given choice of inputs and any given marginal distributions for RAs, expected output will be larger if the RAs tend to be more positively covarying than is the case under independence. The condition is not particularly strong because the PQD dominance need not occur for $\forall \varepsilon \in \mathbb{R}_+^N$, but only for a one-dimensional subset of this; specifically, along $y = \lambda = \varepsilon_i \; \forall \; i \in \Omega_N$. The result should be intuitive in that if there is to be heterogeneity among RAs then it should be as unidimensional as possible in light of the weakest-link constraints.

One final point on RA heterogeneity can be made by considering location shifts in the distribution. Suppose $\varepsilon_i \equiv \mu_i + \eta_i$ where the η_i are independent. Given (1), expected yield may then be represented as

(8) $$E[y \mid x] = \int_{\eta \in \mathbb{R}^N} \min[\mu_1 + \eta_1, \ldots, \mu_N + \eta_N] \prod_{i \in \Omega_N} dF^i(\eta_i \mid x).$$

From (2) and (4), (8) may be alternatively written as

(9) $$E[y \mid x] = \int_{\eta \in \mathbb{R}^N} \text{Prob}[\eta_1 > y - \mu_1, \ldots, \eta_N > y - \mu_N \mid x] dy = \int_0^{y^u} \prod_{i \in \Omega_N} \bar{F}^i(y - \mu_i \mid x) dy.$$

Two definitions are useful at this juncture.

DEFINITION 2: (Marshall and Olkin 1979, p. 7) $Q' \in \mathbb{R}^N$ *is said to be majorized by* $Q'' \in \mathbb{R}^N$*, denoted as* $Q'' \succ Q'$*, whenever both* $\sum_{i=1}^{k} q'_{(i)} \geq \sum_{i=1}^{k} q''_{(i)}\ \forall k \in \Omega_N$ *and* $\sum_{i=1}^{N} q'_{(i)} = \sum_{i=1}^{N} q''_{(i)}$ *where the* $q_{(i)}$ *are the order statistics, i.e.,* $q_{(1)} \leq q_{(2)} \leq \ldots \leq q_{(N)}$*. A function* $W(Q): \mathbb{R}^N \to \mathbb{R}$ *is said to be Schur-concave whenever* $Q'' \succ Q'$ *implies* $W(Q') \geq W(Q'')$*, and it is said to be Schur-convex whenever* $Q'' \succ Q'$ *implies* $W(Q') \leq W(Q'')$.

DEFINITION 3: (Shaked and Shanthikumar 2007, p. 1) *A distribution* $J(\omega): \mathbb{R} \to [0,1]$ *is said to be increasing failure rate (IFR) if* $\mathrm{Ln}[\bar{J}(\omega)]$ *is concave in* ω *while it is decreasing failure rate (DFR) if* $\mathrm{Ln}[\bar{J}(\omega)]$ *is convex in* ω.

Definition 2 captures the idea of more dispersion. To see this, suppose that $\mu'' = \{1,2,6\}$ and $\mu' = \{2,3,4\}$. Then $\mu'' \succ \mu'$ as $2 \geq 1$, $2+3 \geq 1+2$, and $2+3+4 = 1+2+6$. Majorization has been used widely in the economics of income and wealth inequality since the work by Lorenz and Dalton a century ago (Marshall and Olkin 1979, p. 6). Definition 3 seeks to measure how quickly a distribution tail tapers off, where IFR identifies a rapidly fading right-hand tail. Our interest in majorization is when the μ_i location parameters become more dispersed in the sense of a majorization shift. A rapidly fading right tail for each marginal RA distribution suggests that dispersion in location shifts for independent draws from otherwise identical distributions will reduce the expected value of the least-order statistic and so will reduce expected yield. The next proposition confirms this.

PROPOSITION 3: *Suppose the technology is LoM, while input-conditioned RAs are independent* *and have a common distribution up to location. Let the distribution express IFR (DFR). Then a majorizing shift in the location vector reduces (increases) expected output for any given input choice.*

The proposition gives precise conditions under which heterogeneity in the technology of RAs is detrimental to anticipated yield for any given set of inputs. Perhaps contrary to the intuition one might glean from Liebig's barrel, eqn. (1), and Proposition 1, even under the very stylized setting of Proposition 3 we find that heterogeneity in RAs need not adversely impact yield. Bear in mind though that for a distribution function to be DFR at a point, the density function must be decreasing at that point. So for DFR to apply over the entire support, it must be that the density function is decreasing over the support.

What sorts of distributions exhibit DFR? A commonly used distribution in reliability theory, which is what our study of crop yield distributions has brought us to, is the Weibull. It has the form

(10) $$F^i(\varepsilon_i \mid x) = 1 - e^{-[\lambda(x)\varepsilon_i]^\alpha},$$

for $\varepsilon_i \geq 0$, $\lambda(x) > 0$, and $\alpha > 0$ (Rausand and Høyland 2004). As is readily shown, the distribution expresses IFR if $\alpha \geq 1$ and DFR if $\alpha \leq 1$. Our interest is in the product of location-displaced survival functions along the equal values line. Using $\text{Prob}(\varepsilon_i + \mu_i > y) =$ $\text{Prob}(\varepsilon_i > y - \mu_i)$ and location-shifted univariate Weibull survival functions, then the yield survival function under independence is

(11) $$\prod_{j=1}^{N} e^{-[\lambda(x)]^\alpha (y-\mu_i)^\alpha} = e^{-[\lambda(x)]^\alpha \sum_{i=1}^{N} (y-\mu_i)^\alpha}.$$

It can readily be shown that this is smaller under a more dispersed location vector when $\sum_{i=1}^{N}(y-\mu_i)^{\alpha}$ is Schur-convex, and that occurs when $\alpha > 1$. On the other hand, (11) increases with more dispersion in the location vector when $\alpha < 1$, and there is no effect when $\alpha = 1$.

An alternative model of resource availabilities is the gamma with location displacements. Here,

$$(12) \qquad F^{i}(\varepsilon_i \mid x) = \frac{[\lambda(x)]^{\alpha}}{\Gamma(\alpha)} \int_{s\in(0,\varepsilon_i]} s^{\alpha-1} e^{-\lambda(x)s} ds,$$

for $\lambda(x) > 0$, $\alpha > 0$, and $\Gamma(\alpha)$ the gamma function. The yield survival function is given by

$$(13) \qquad [\lambda(x)]^{N\alpha} \left[\Gamma(\alpha)\right]^{-N} \prod_{i=1}^{N} \int_{y-\mu_i}^{\infty} s^{\alpha-1} e^{-\lambda(x)s} ds.$$

It is well known that the gamma distribution is IFR if $\alpha \geq 1$ and DFR if $\alpha \leq 1$ (Rausand and Høyland 2004, p. 61). Use of Definition 3 and a little further work shows that the yield survival function (and so expected yield) is decreasing with more dispersion in the location shifters whenever $\alpha \geq 1$ and increasing whenever $\alpha \leq 1$.[3]

Comparing Propositions 2 and 3, one sees that care is required when stylizing heterogeneity in a crop's technology. Given marginals, then less heterogeneity in the sense of more positively covarying RAs is good. Given dependence structure, namely independence, then more heterogeneity in mean is probably bad, but we cannot be sure without further knowledge on the marginal distributions. Tail thickness matters. We turn next to the issue of skewness, where RA tail thickness will assume a more prominent role.

---

[3] Apply the Ostrowski method from the proof of Proposition 3, in the appendix. From an empirical perspective, Babcock and Blackmer (1992) have modeled soil nitrate availability for Iowa corn production using a location shifted gamma distribution in a LoM technology framework. They find $\alpha \geq 1$; see their Table 1.

Skewness and Uniform Case

The yield distribution of interest is when inputs are fixed, for otherwise the measured statistical attributes of yield may be due to heterogeneity in input uses over the area of interest and not due to the technology itself. To make further progress in this section, we also assume that each ε_i is independently drawn from the standard continuous uniform distribution, $\varepsilon_i \sim U[0,1]$. Thus there is no bias in favor of any resource, and in addition the RAs are symmetric around the mean so that they have zero skewness. A well-known result is that the first-order statistic then has density

$$(14) \qquad g(y;1,N) = \frac{(1-y)^{N-1}}{\int_0^1 (1-z)^{N-1} dz},$$

or the beta distribution with parameters $\alpha = 1$ and $\beta = N$ (Gupta and Nadarajah 2004a, p. 89). Its mean, variance, and third central moment are $\mu_y = (1+N)^{-1}$, $\sigma_y^2 = N(1+N)^{-2}(2+N)^{-1}$, and $\zeta_y = 2N(N-1)(2+N)^{-1}(3+N)^{-1}(1+N)^{-3} > 0$. Bearing in mind that skewness is defined as $\gamma_y = \zeta_y / \sigma_y^3$, we have

$$(15) \qquad \gamma_y = \frac{2(N-1)}{(3+N)} \sqrt{\frac{2+N}{N}} > 0.$$

Thus, yield in this case expresses positive skew as found in Day (1965) for Mississippi cotton and corn during 1921-57 and Ramirez, Misra, and Field (2003) for West Texas Dryland cotton during 1970-99.

Heterogeneity in Means

Of course, in reality even if RAs have uniform marginals, they are unlikely to have common means or variances, if only because factor prices, agronomic knowledge, and technological capabilities differ. Neither are the RAs likely to be independent. We will relax each of these conditions in turn. To focus on effects, let there be just two resources at issue, where ε_1 is discrete uniformly distributed on point pair $\{0,1\}$, ε_2 is discrete uniformly distributed on $\{\tau, 1+\tau\}$, $\tau \in [0,1]$, and these random variables are independent.[4] Therefore, the random variables have the same higher central moments for marginals, differing only by the shifted mean. In order to commence with a zero skew distribution, let the probability of both low and high states be 0.5. The distribution of $y = \min[\varepsilon_1, \varepsilon_2]$ is 0 with probability 0.5, τ with probability 0.25, and 1 with probability 0.25. The moments are $\mu_y = (\tau+1)/4$, $\sigma_y^2 = (3\tau^2 - 2\tau + 3)/16$, $\zeta_y =$ $3(1-\tau)^2(1+\tau)/32$ with skewness $\gamma_y = 6(1-\tau)^2(1+\tau)(3\tau^2 - 2\tau + 3)^{-3/2} > 0$ where the derivative of interest satisfies

(16) $$\frac{d\gamma_y}{d\tau} = -\frac{96(1-\tau)\tau}{\left(3\tau^2 - 2\tau + 3\right)^{2.5}} < 0$$

on $\tau \in [0,1]$. In addition, for $\tau \in [-1,0]$, eqn. (16) shows that an increase in τ leads to a less negative skewness statistic. Thus, heterogeneity in location alone tends to reduce skewness for the discrete uniform distribution. This is because the location shift takes probability mass away from a support point at the lower end of the distribution.

Heterogeneity in Variances

As above, let there be just two resources at issue, where ε_1 is discrete uniformly distributed on $\{0,1\}$. But let ε_2 be discrete uniformly distributed on $\{\tau, 1-\tau\}$, $\tau \in [0, 0.5]$, while these random

---

[4] If $\tau > 1$ then $\min[\varepsilon_1, \varepsilon_2] \equiv \varepsilon_1$.

variables are (again) independent. If $\tau = 0.5$ then the support of ε_2 is concentrated at $\varepsilon_2 = 0.5$. As before, set the skew as zero by letting the probability of the low state be 0.5 in each case. The distribution of $y = \min[\varepsilon_1, \varepsilon_2]$ is 0 with probability 0.5, τ with probability 0.25, and $1-\tau$ with probability 0.25. The moments are $\mu_y = 0.25$, $\sigma_y^2 = (\tau^2 - 8\tau + 3)/16$, and $\zeta_y = 3(1-2\tau)^2/32$ with skewness $\gamma_y = 6(1-2\tau)^2(8\tau^2 - 8\tau + 3)^{-3/2}$ so that the derivative of interest is

$$(17)\qquad \frac{d\gamma_y}{d\tau} = \frac{96(2\tau-1)\tau(1-\tau)}{\left(3-8\tau+8\tau^2\right)^{2.5}} \overset{sign}{=} \tau - 0.5 < 0.$$

Heterogeneity in variance, through contracting the support of one distribution, reduces skewness. As with a location shift, heterogeneity takes probability mass away from a heavily weighted support point in the distribution's left tail.

Dependence

Following Dasgupta and Maskin (1987), set

$$(18)\qquad (\varepsilon_1, \varepsilon_2) = \begin{cases} (0,0) & \text{with probability} \quad 0.25(1+\rho); \\ (0,1) & \text{with probability} \quad 0.25(1-\rho); \\ (1,0) & \text{with probability} \quad 0.25(1-\rho); \\ (1,1) & \text{with probability} \quad 0.25(1+\rho); \end{cases}$$

for $\rho \in [-1,1]$ where $\rho > 0$ acts to place more probability on points $(0,0)$ and $(1,1)$ so this is an illustration of a probability shift given in Definition 1. Our interest here is not in understanding the impact on mean but rather on higher moments. Mean, variance, and third central moment of yield are now $\mu_y = 0.25(1+\rho)$, $\sigma_y^2 = (3-\rho)(1+\rho)/16$, $\zeta_y = (3-\rho)(1+\rho)(1-\rho)/32$, where skewness is $\gamma_y = 2(1-\rho)(3-\rho)^{-0.5}(1+\rho)^{-0.5}$. Note that $\mathrm{Lim}_{\rho \downarrow -1}\gamma_y = +\infty$ and $\gamma_y = 0$ when $\rho = 1$. The derivative with respect to the correlation parameter is

(19) $$\frac{d\gamma_y}{d\rho} = -\frac{8}{(3-\rho)^{3/2}(1+\rho)^{3/2}} < 0.$$

So an increase in correlation decreases skewness for the discrete uniform distribution. As correlation increases, the distribution of $\min[\varepsilon_1, \varepsilon_2]$ becomes more like the uniform distribution, with zero skewness. A decrease in correlation can be seen as an increase in heterogeneity in the RAs. Overall, we see that skewness falls when means diverge (more heterogeneity), one of the random variables assumes less variance (more heterogeneity), or the distributions become more correlated (less heterogeneity).

Skewness and Normal Case

Now let the ε_i be independent and drawn from an identical distribution, namely, $F^i(\varepsilon_i) = \Phi\big((\varepsilon_i - \mu)/\sigma\big)$, the standard cumulative normal before being relocated and scaled by common parameters. Then the yield density for the minimum of N draws is

(20) $$g(y) = \frac{N}{\sigma}\left[1-\Phi\left(\frac{y-\mu}{\sigma}\right)\right]^{N-1}\phi\left(\frac{y-\mu}{\sigma}\right),$$

where $\phi(\cdot)$ is the density of the normalized distribution. This is an instance of the Beta-Normal distribution as discussed in Eugene, Lee, and Famoye (2002).[5] When $N = 2$, then Choi (2005), correcting Gupta and Nadarajah (2004b), shows the first three moments to be $\mu_y = \mu - \sigma\pi^{-0.5} < \mu$, $\sigma_y^2 = (\pi-1)\sigma^2\pi^{-1}$, and $\zeta_y = 0.5(\pi-4)\sigma^3\pi^{-3/2} < 0$. Thus, and by contrast with the uniform

[5] Actually, as has been recently pointed out by Jones (2004), when N is a natural number, then this is the least-order statistic of independent draws from a common normal distribution. The literature on order statistics for i.i.d. normal random variables has a long pedigree dating at least as far back as Bose and Gupta (1959).

distribution, skewness is negative at $\gamma_y = 0.5(\pi - 4)(\pi - 1)^{-3/2}$. This should not be surprising when one considers how weighting $1 - \Phi((y-\mu)/\sigma)$ biases density $\phi((y-\mu)/\sigma)$ in (20). Though low at low yields, the yield density function should not be as low for low yield draws as for high yield draws. This is in contrast to uniform RA densities, where the survival function density in (14) completely determines the shape of yield density, and yield density will have negative derivative everywhere it has support.

To illustrate the effects of heterogeneity, suppose first that $N = 2$ and the random variables are perfectly positively correlated. Then yield follows the standard normal up to location and scale, thus having zero skew. Suppose instead that the random variable had perfect negative correlation so that $(\varepsilon_1 - \mu)/\sigma \equiv -(\varepsilon_2 - \mu)/\sigma$. Then the yield distribution is half-normal with support to the left of μ and clearly has negative skew. By contrast with the uniform, an increase in correlation (and so a decrease in heterogeneity) from $\rho = 0$ to $\rho = 1$ increases skewness to $\gamma_y = 0$.

The case of drawing from a bivariate normal with heterogeneous means involves a nonlinearity. If, under independence and equal variances, means differ substantially, then only the marginal with the lower mean matters, and skewness should not differ much from zero. With means $\bar{\varepsilon}_1$ and $\bar{\varepsilon}_2$, as $\bar{\varepsilon}_1$ increases toward $\bar{\varepsilon}_2$ then yield skewness should decrease away from zero until $\bar{\varepsilon}_1 = \bar{\varepsilon}_2$ and should increase toward zero thereafter. So heterogeneity in means should increase skewness toward zero. This is in contrast to the uniform case where heterogeneity in means decreases skewness toward zero.

Alternatively, suppose we allow possibly distinct means and variances on the RAs where the variance of ε_2 recedes to zero. Then, of course, correlation ceases to be a meaningful statistic and we may ignore it. With mean of ε_2 at $\bar{\varepsilon}_2$, the distribution becomes censored normal with support to the left of $\bar{\varepsilon}_2$ and strictly positive probability massed at $\bar{\varepsilon}_2$. Again, the distribution is clearly negatively skewed. So casual heuristic reasoning suggests that heterogeneity through lower variance for one RA tends to introduce a downward bias in the skewness of the least-order statistic of symmetric random variables when compared with the baseline case where independent draws are taken from identical distributions. This downward bias is as under the comparable scenario for the uniform distribution.

Unfortunately, thought exercises such as the above aside, little appears to be known about order statistics for draws from non–i.i.d. normal distributions. To investigate further, Table 1 provides moment estimates for a variety of scenarios when $N = 2$. The benchmark is italicized and bolded with means at $\bar{\varepsilon}_1 = \bar{\varepsilon}_2 = 10$, standard deviations at $\sigma_1 = \sigma_2 = 1$ where $\sigma_i^2 = \mathrm{Var}(\varepsilon_i)$, and correlation at $\rho = 0$. After taking 10,000 independent draws, antithetic variates were used to double the sample to 20,000.[6]

Note first that the mean and standard deviation of yield increase with an increase in correlation. For yield mean, per Definition 1 and Proposition 2, it is preferable that low draws come in pairs in order to get them out of the way. For standard deviation, the LoM minimization operation pushes probability weightings toward the lower end of the support so that one should expect dispersion to decline. Notice also that skewness is never positive, i.e., the simulations

---

[6] For antithetic variates, if the draw $(\varepsilon_1, \varepsilon_2) = (\hat{\varepsilon}_1, \hat{\varepsilon}_2)$ is made then so is $(\varepsilon_1, \varepsilon_2) = (-\hat{\varepsilon}_1, -\hat{\varepsilon}_2)$. The intent is to promote law-of-large-numbers convergence by balanced sampling (Boyle, Broadie,

suggest that what applies for the first-order statistic with i.i.d. normal draws may be robust to relaxing the i.i.d. requirements.

Confining attention now to zero correlation simulations, yield skewness does not differ much across differences in means when variances are common. But skewness is more strongly negative when the common variance is large. Heterogeneity in means only ensures that the distribution with the lower mean dominates when determining the first-order statistic. With a sufficiently large gap in means and sufficiently low standard deviations, the yield skew will be close to zero.

For $\bar{\varepsilon}_1 = \bar{\varepsilon}_2 = 10$ too, and recalling that skewness is normalized by yield variance, yield skewness is more negative when there is also heterogeneity in the RA variances. This is as with the uniform distribution but for a somewhat different reason. There a contraction in the variance of one RA ensured that the upper tail of the yield distribution was increasingly concentrated around the mean of that RA. An increase in the RA variance would not have that effect. For normally distributed RAs with equal means, a decrease in the variance of one tends to induce a more negative skew because one is converging on a censored (from the top) normal yield distribution. An increase in the variance of one RA also tends to induce a more negative skew. This is because the more dispersed distribution dominates in determining the left end of yield density whereas the tightly dispersed distribution dominates in determining the right end of yield density.

The case $(\rho, \sigma_1, \sigma_2) = (0,1,2)$ merits attention. There, skewness is -0.664 under $\bar{\varepsilon}_1 = \bar{\varepsilon}_2 = 10$, -0.246 under $(\bar{\varepsilon}_1, \bar{\varepsilon}_2) = (8,10)$ and -0.448 under $(\bar{\varepsilon}_1, \bar{\varepsilon}_2) = (12,10)$. Skewness becomes less negative when $\bar{\varepsilon}_1$ decreases from 10 to 8 and when $\bar{\varepsilon}_1$ increases from 10 to 12. At $\bar{\varepsilon}_1 = 8$ then ε_1

---

and Glasserman 1997).

is dominant in determining yield because it has lower mean and lower variance. So skewness should move toward 0, that of the univariate normal distribution. The move toward 0 skewness is not as pronounced when $\bar{\varepsilon}_1$ changes from 10 to 12 because the large variance of ε_2 ensures it can still throw up a good draw such that ε_1 narrowly distributed at the upper end of the yield range will still be important.

Moving to positive correlation, with $\bar{\varepsilon}_1 = \bar{\varepsilon}_2 = 10$ then the effect on skewness is clear. When compared with $\rho = 0$, skewness is always less negative. Negative correlation makes skewness even more negative for any given set of RA mean and standard deviation parameters. The most negative skew statistic arises when $(\rho, \bar{\varepsilon}_1, \bar{\varepsilon}_2, \sigma_1, \sigma_2) = (-0.5, 10, 10, 1, 0.5)$, followed closely by $(\rho, \bar{\varepsilon}_1, \bar{\varepsilon}_2, \sigma_1, \sigma_2) = (-0.5, 10, 10, 1, 2)$. Here, the common mean ensures that both marginals are relevant. The negative correlation ensures that a moderately below-average draw from one marginal is very likely to matter, thus facilitating a pileup of probability mass toward the RA means. Variance heterogeneity allows for a thinly spread out left tail to the yield distribution.

Skewness and Raised Cosine Case

To further probe the conjecture that the distribution tail determines skewness, consider the raised cosine distribution. In this case, let $y = \min[\varepsilon_1, \varepsilon_2]$ where the ε_i are independent with common density $f(\varepsilon_i) = 0.5\mathrm{Cos}(\varepsilon_i)$ on $[-\pi/2, \pi/2]$, the cosine function's domain of positive value.[7] The

[7] Here, 0.5 normalizes since $\int_{-\pi/2}^{\pi/2} \mathrm{Cos}(\varepsilon_i) d\varepsilon_i = \mathrm{Sin}(\varepsilon_i)\big|_{-\pi/2}^{\pi/2} = 2$. The support was chosen for analytic convenience. While it intersects the negative domain, remember that skewness is location independent. A location shift of support $[-\pi/2, \pi/2] \to [\tau - \pi/2, \tau + \pi/2]$, $\tau \geq -\pi/2$, would not affect skewness.

distribution is chosen because the power transformation $0.5\text{Cos}(\varepsilon_i) \rightarrow \kappa(\alpha)[\text{Cos}(\varepsilon_i)]^\alpha$ thins out the density tails when $\alpha > 1$, and because the class of densities is analytically tractable. The cumulative distribution for RA density $f(\varepsilon_i) = 0.5\text{Cos}(\varepsilon_i)$ is $F(\varepsilon_i) = [1 + \text{Sin}(\varepsilon_i)]/2$. From (3),

$$(21) \qquad g(y) = \frac{\text{Cos}(y)[1 - \text{Sin}(y)]}{2},$$

while moments are $\mu_y = -0.3927$, $E[y^2] = 0.4674$, $E[y^3] = -0.3799$, and $\gamma_y = E[(y - \mu_y)^3]/\sigma_y^3$ $= 0.28$.[8] By contrasting (20) with (21), it can be seen how tail thickness ensures positive skewness for the first-order statistic in this case.

Instead, suppose the ε_i are independent with common density $(0.5\pi)^{-1}\text{Cos}^2(\varepsilon_i)$ on $[-\pi/2, \pi/2]$, the cosine function's domain of positive value.[9] Squaring the smaller values close to the support boundaries thins out the tails. The cumulative distribution for a resource availability is $0.5 + [\varepsilon_i + \text{Cos}(\varepsilon_i)\text{Sin}(\varepsilon_i)]/\pi$. From (3), yield density is

$$(22) \qquad g(y) = \frac{2\text{Cos}^2(y)\left\{0.5 - ([y + \text{Cos}(y)\text{Sin}(y)])/\pi\right\}}{\pi},$$

while moments are $\mu_y = -0.1623$, $E[y^2] = 0.1612$, $E[y^3] = -0.1150$, $\gamma_y = -0.9086$. Thinning out the tails of the RA distribution changes the yield skew from positive to negative.

Extreme Value Analysis

It was mentioned in Just and Weninger (1999) that crop yield statistics, being averages over space and perhaps over time too, should comply with a relevant central limit theorem as the

---

[8] Integrations were performed with the assistance of the Wolfram Integrator webMATHEMATICA, available at http://integrals.wolfram.com/index.jsp.

limiting distribution. Bear in mind that the limiting distribution for the average of independently drawn random variables (independence requirement) from a common distribution (homogeneity requirement) is just the distribution mean with zero values for all centered higher-order moments. This is due to the strong law of large numbers under mild regularity conditions (Durrett 1996, p. 56). Central limit theorems convey the way in which convergence to this distribution mean occurs, and scaling by root sample size $N^{0.5}$ is necessary to avoid a degenerate limiting distribution.

While accepting that central limit theorems are relevant, our intent is to set aside aggregation issues by considering a sufficiently small and homogeneous area so that all relevant stochastic realizations and consequences are the same. Returning to (1), and primarily as a theoretical counterpoint to the Just and Weninger argument, assume the ε_i are independently, identically drawn while N is large. Yield being the first-order statistic, we are now not in the realm of limiting distributions for arithmetic averages but rather in that of limiting distributions for extreme order statistics (Coles 2001; de Haan and Ferreira 2006; Bain and Engelhardt 1992).[10]

The relevant distribution for convergence is that of the generalized extreme value distribution (Coles 2001, p. 47; de Haan and Ferreira 2006, p. 6).[11] If it exists, then the limiting distribution *for the minimum* is, in general form, that of Von Mises and is called the Generalized Extreme

---

[9] The normalizing constant is 0.5π since $\int_{-\pi/2}^{\pi/2} \mathrm{Cos}^2(\varepsilon_i) d\varepsilon_i = 0.5[\varepsilon_i + \mathrm{Cos}(\varepsilon_i)\mathrm{Sin}(\varepsilon_i)]\Big|_{-\pi/2}^{\pi/2} = 0.5\pi$.

[10] A theory of central limits for statistical aggregates that encompasses both averages (the usual case) and extreme values has been developed; see Schlather (2001) and Bogachev (2006).

[11] As with the central limit theorem for averages, the limiting distribution of the minimum for independent draws from a common distribution is trivially degenerate. The distribution at issue is for N large but not too large.

Value distribution (Castillo 1988, p. 108; Coles 2001, p. 53):[12]

$$(23)\qquad J(y) = 1 - e^{\left\{-\left[1+c\left(\frac{\lambda - y}{\delta}\right)^{-1/c}\right]\right\}}; \qquad \delta > 0; \qquad 1 + \left(\frac{\lambda - y}{\delta}\right)c \geq 0.$$

In this three-parameter family, λ and δ may be viewed (loosely) as location and scale parameters while c determines shape. The minimum of independent identical draws cannot have limiting distribution other than this form, just as the normal can be the only limiting distribution for averages. Its attractive property is replicability or min-stability whereby the minimum of independent draws from the distribution follows the same distribution up to location and scale.[13]

This distribution can take one of three specific forms, depending on the value of parameter c. These are Fréchet for $c > 0$, Weibull for $c < 0$, and Gumbel for $c = 0$, where the convergence limit in (23) as $c \to 0$ is (Castillo 1988, p. 185)

$$(24)\qquad \text{Gumbel}: \quad J(y) = 1 - e^{-e^{-\left(\frac{\lambda - y}{\delta}\right)}}, \qquad y \in \mathbb{R},$$

with mean value $\lambda - 0.5777\delta$, median $\lambda - 0.3665\delta$, variance $\pi^2\delta^2/6$, and skewness -1.1396.

Each of these specific distributions has a domain of attraction, i.e., a distribution function domain such that the first-order statistic of a set of independent draws converges to this form. The case of $c > 0$ is ruled out from consideration because the minimum of draws from a distribution with finite lower bound, as with resource availabilities, cannot converge to the Fréchet distribution (Castillo 1988, p. 102). Upon considering (3), (14), (20), and (21), it should be no surprise that the determinant of which form, if any, a given distribution is attracted to is tail

[12] As we will see below, the similarity of the exponent with the HARA utility specification is not incidental.

[13] The literature generally refers to the counterpart for the maximum extremum, max-stability. But $\max[\varepsilon_1, \ldots, \varepsilon_N] \equiv -\min[-\varepsilon_1, \ldots, -\varepsilon_N]$, ensuring the solutions are almost identical.

behavior in the tail of interest (Castillo 1988, pp. 100-120; de Haan and Ferriera 2006, pp. 33). Thick tails, as with the uniform distribution, should be expected to behave differently when compared with thin tail distributions such as the normal.

Specifically, a necessary and sufficient condition for convergence to the Gumbel as limiting distribution is that (Castillo 1988, p. 102)

(25) $$\mathrm{Lim}_{t\downarrow 0} \frac{F\left(F^{-1}(t)+\left[F^{-1}(t/e)-F^{-1}(t)\right]z\right)}{t}=e^{-z}.$$

When $z=0$, then $F[F^{-1}(t)]/t$ does equal 1. Notice too that $F^{-1}(t)+[F^{-1}(t/e)-F^{-1}(t)]z$ may be viewed as a directed first-order Taylor series expansion of quantile function $F^{-1}(\cdot)$ around t. Since $t/e \approx t/2.7183 < t$, it follows that $F^{-1}(t/e) \le F^{-1}(t)$, the direction of the expansion is to the left of the t-quantile, and the behavior at issue is on the left tail. Turning to the Weibull as limiting distribution, a necessary and sufficient condition for convergence is that (Castillo 1988, p. 114)

(26) $$\mathrm{Lim}_{t\downarrow 0} \frac{F^{-1}(2t)-F^{-1}(t)}{F^{-1}(4t)-F^{-1}(2t)}<1,$$

where the limit is required to exist. If a distribution has a thick left tail then $F(\cdot)$ rises sharply near its lower support so that $F^{-1}(2t)-F^{-1}(t)$ is likely to be small when compared with quantile differences a little further from the left support.

For minima, the normal, lognormal and gamma distributions have the Gumbel distribution given in (24) as limiting distribution whereas the uniform and exponential have Weibull as the limiting distribution (Castillo 1988, p. 120). The Weibull distribution considered in (23) is not that usually studied, as in Bain and Engelhardt (1992). Rather, it is the mirror image up to relocation from the origin.[14] By contrast with the Gumbel distribution for minima, the standard Weibull distribution can have positive or negative skewness. So knowledge that the first-order statistic of the uniform distribution has Weibull distribution as limiting distribution leaves us no wiser without further information. Thus, there is some evidence to believe that a sufficiently thin left tail on RA distributions will tend to support a negatively skewed crop yield distribution whenever that distribution is determined by a LoM technology.

Conclusion

This article has used the law of the minimum, or weakest-link crop production technology, together with structure on the input-conditioned resource availabilities to seek a better understanding of the stochastic attributes of crop yield distributions. Some curiosities have been identified. For instance, when each applied input is matched to just one resource availability and first-order dominating shifts are induced in the marginal, then negative correlation among resource availabilities can never overturn the tendency for inputs to complement under LoM. The role of stochastic dependence structures was investigated to provide precise conditions under which positive dependency between given marginals for resource availabilities will increase expected yield when compared with independence. In addition, it was shown that conditions exist under which location-shifting heterogeneities in resource availabilities can increase expected yield. This observation is perhaps surprising in light of the Liebig barrel analogy that so beautifully characterizes the LoM under certain resource availabilities.

Yield skewness was considered for three types of resource availability distributions. It was

[14] For the maximum-order statistic, the relocated Weibull with usual orientation is the appropriate limiting distribution.

shown that the LoM can support both positive and negative yield skewness. Location and scale heterogeneities were studied to discern definite, but sometimes involved, patterns in their implications for skewness. It was suggested that the left tail attributes of resource availability distributions are key in determining yield skew, and a connection with extreme value theory was provided. Again, this theory can support either positive or negative skewness for zero skew, independent and identical resource availability distributions. If the crop production process is quite tightly controlled, then the left tails of resource availabilities should be thin, and negative skewness will be favored. This suggests that one should be more likely to compute negative skewness when looking at yield data of more recent vintage, in prime growing areas, and in more developed countries where market inputs are more readily available.

References

Babcock, B.A., and A. M. Blackmer. 1992. "The Value of Reducing Temporal Input Nonuniformities." *Journal of Agricultural and Resource Economics* 17(December):335–347.

Bain, L.J., and M. Engelhardt. 1992. *Introduction to Probability and Mathematical Statistics*, 2nd ed. Belmont, CA: Duxbury Press.

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

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

Bloom, A.J., F.S. Chapin, III, and H.A. Mooney. 1985. "Resource Limitations in Plants–An Economic Analogy." *Annual Review of Ecology and Systematics* 16:363–392.

Bogachev, L. 2006. "Limit Laws for Norms of IID Samples with Weibull Tails." *Journal of Theoretical Probability* 19(October):849–873.

Bose, R.C., and S.S. Gupta. 1959. "Moments of Order Statistics from a Normal Population." *Biometrika* 46(December):433–440.

Boyle, P.P., M. Broadie, and P. Glasserman. 1997. "Monte Carlo Methods for Security Pricing." *Journal of Economic Dynamics and Control* 21(June):1267–1321.

Castillo, E. 1988. *Extreme Value Theory in Engineering*. Boston: Academic Press.

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.

Chambers, R.G., and E. Lichtenberg. 1996. "A Nonparametric Approach to the von Liebig-Paris

Technology." *American Journal of Agricultural Economics* 78(May):373–386.

Chapin, F.S., III, A. Bloom, C. Field, and R. Waring. 1987. "Plant Responses to Multiple Environmental Factors." *Bioscience* 37(January):49–57.

Choi, J.-K. 2005. "Letter to the Editor." *Communications in Statistics–Theory and Methods* 34(4):997–998.

Coles, S. 2001. *An Introduction of Statistical Modeling of Extreme Values*. Springer.

Cornes, R. 1993. "Dike Maintenance and Other Stories: Some Neglected Types of Public Goods." *Quarterly Journal of Economics* 108(February):259–271.

Dasgupta, P., and E. Maskin. 1987. "The Simple Economics of Research Portfolios." *Economic Journal* 97(September):581–595.

David, H.A., and H.N. Nagaraja. 2003. *Order Statistics*. Hoboken, NJ: Wiley.

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

de Haan, L., and A. Ferreira. 2006. *Extreme Value Theory: An Introduction*. New York: Springer.

Dow, W.H., T.J. Philipson, and X. Sala-I-Martin. 1999. "Longevity Complementarities under Competing Risks." *American Economic Review* 89(December):1358–1371.

Durrett, R. 1996. *Probability: Theory and Examples*, 2nd ed. Belmont, CA: Duxbury Press.

Eugene, N., C. Lee, and F. Famoye. 2002. "Beta-Normal Distribution and Its Applications." *Communications in Statistics–Theory and Methods* 31(4):497–512.

Frank, M.D., B.R. Beattie, and M.E. Embleton. 1990. "A Comparison of Alternative Crop Response Models." *American Journal of Agricultural Economics* 72(August):597–603.

Gallagher, P. 1987. "U.S. Soybean Yields: Estimation and Forecasting with Nonsymmetric

Disturbances." *American Journal of Agricultural Economics* 69(November):796–803.

Gupta, A.K., and S. Nadarajah. 2004a. "Order Statistics and Records." In A.K. Gupta and S. Nadarajah, eds. *Handbook of Beta Distribution and Its Applications*. New York: Marcel Dekker, Inc., pp. 88–96.

Gupta, A.K., and S. Nadarajah. 2004b. "On the Moments of the Beta-Normal Distribution." *Communications in Statistics–Theory and Methods* 33(1):1–13.

Hirshleifer, J. 1983. "From Weakest-Link to Best-Shot: The Voluntary Provision of Public Goods." *Public Choice* 41(June):371–386.

Jones, M.C. 2004. "Letter to the Editor: The Moments of the Beta-Normal Distribution with Integer Parameters are the Moments of Order Statistics from the Normal Distribution." *Communications in Statistics–Theory and Methods* 33(11):2869–2870.

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

Lehmann E.L. 1966. "Some Concepts of Dependence." *Annals of Mathematical Statistics* 37(October):1137–1153.

Llewelyn, R.V., and A.M. Featherstone. 1997. "A Comparison of Crop Production Functions Using Simulated Data for Irrigated Corn in Western Kansas." *Agricultural Systems* 54(August):521–538.

Lynch, J.P., and M.D. Ho. 2005. "Rhizoeconomics: Carbon Costs of Phosphorus Acquisition." *Plant and Soil* 269(February):45–56.

Marshall, A.W., and I. Olkin. 1979. *Inequalities: Theory of Majorization and Its Applications*. San Diego: Academic Press.

Milgrom, P., and C. Shannon. 1994. "Monotone Comparative Statics." *Econometrica*

62(January):157–180.

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

Nelsen, R.B. 1999. *An Introduction to Copulas*. Lecture Notes in Statistics, #139. New York: Springer.

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

Nelson, C.H. 1990. "The Influence of Distributional Assumptions on the Calculation of Crop Insurance Premia." *North Central Journal of Agricultural Economics* 12(January):71–78.

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

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

Rastetter, E., and G. Shaver. 1992. "A Model of Multiple Limitation for Acclimating Vegetation." *Ecology* 73(August):1157–1174.

Rausand, M., and A. Høyland. 2004. *System Reliability Theory: Models, Statistical Methods, and Applications*, 2nd edn. Hoboken NJ: Wiley.

Rubio, G., J. Zhu, and J.P. Lynch. 2003. "A Critical Test of the Two Prevailing Theories of Plant Response to Nutrient Availability." *American Journal of Botany* 90(January):143–152.

Schlather, M. 2001. "Limit Distributions of Norms of Vectors of Positive i.i.d. Random Variables." *Annals of Probability* 29(April):862–881.

Shaked, M., and J.G. Shanthikumar. 2007. *Stochastic Orders*. New York: Springer.

Sherrick, B.J., F.C. Zanini, G.D. Schnitkey, and S.H. Irwin. 2004. "Crop Insurance Valuation under Alternative Yield Distributions." *American Journal of Agricultural Economics* 71(May):406–419.

Sinclair, T.R., and W.I. Park. 1993. "Inadequacy of the Liebig Limiting-Factor Paradigm for Explaining Varying Crop Yield." *Agronomy Journal* 85(May/June):742–746.

Swinton, S.M., and R.P. King. 1991. "Evaluating Robust Regression Techniques for Detrending Crop Yield Data with Nonnormal Errors." *American Journal of Agricultural Economics* 73(May):446–451.

Appendix

Proof of Proposition 1: We will use Sklar's theorem (Nelsen 1999, p. 41), namely (as stated for survival functions), the fact that any multivariate survival function with defined marginal survival functions $\bar{F}^i(\varepsilon_i \mid x)$ can be represented in the copula form of $\hat{C}[\bar{F}^1(\varepsilon_1 \mid x), \ldots, \bar{F}^N(\varepsilon_N \mid x)]$ where the properties that $\hat{C}[\cdot]$: $[0,1]^N \to [0,1]$ must satisfy include 2-increasing. A 2-increasing function $f(z_1, \ldots, z_N)$ is one that satisfies $f(\ldots, z_i'', \ldots, z_j'' \ldots) + f(\ldots, z_i', \ldots, z_j' \ldots) \geq f(\ldots, z_i', \ldots, z_j'' \ldots)$ $+ f(\ldots, z_i'', \ldots, z_j' \ldots) \forall i, j \in \Omega_N, \forall z_i'' \geq z_i', \forall z_j'' \geq z_j'$, i.e., it is supermodular and would have non-negative cross-derivatives were it twice continuously differentiable. From (2), our interest is in $\bar{G}(y \mid x) = \hat{C}[\bar{F}^1(y \mid x), \ldots, \bar{F}^N(y \mid x)]$, or along the copula bisector $y = \varepsilon_i \ \forall i \in \Omega_N$. The cross differentiation with respect to x_i and x_j gives

$$\text{(A1)} \quad \sum_{l=1}^{N} \overbrace{\frac{\partial \hat{C}[u^1, \ldots, u^N]}{\partial u^l}}^{+} \overbrace{\frac{\partial^2 \bar{F}^l(y \mid x)}{\partial x_i \partial x_j}}^{+} + \sum_{r=1}^{N} \sum_{l=1}^{N} \overbrace{\frac{\partial^2 \hat{C}[u^1, \ldots, u^N]}{\partial u^l \partial u^r}}^{+} \overbrace{\frac{\partial \bar{F}^l(y \mid x)}{\partial x_i}}^{+} \overbrace{\frac{\partial \bar{F}^r(y \mid x)}{\partial x_j}}^{+} \geq 0,$$

where $u^k \equiv \bar{F}^k(y \mid x_i)$. Given twice continuous differentiability of $\bar{G}(y \mid x)$ and *i*)-*ii*), (A1) must be true over $y \in [0, y^u]$. That is all we need to prove in light of (2) and (4).

Proof of Proposition 3: Ostrowski's criterion (Marshall and Olkin 1979, p. 57) asserts that a continuously differentiable function $h(z): I_1 \to I_2, I_1 \subset \mathbb{R}^N, I_2 \subset \mathbb{R}$ is Schur-concave whenever

$$\text{(A2)} \quad \left(\frac{\partial h(z)}{\partial z_i} - \frac{\partial h(z)}{\partial z_j} \right)(z_i - z_j) \leq 0,$$

and Schur-convex whenever the inequality in (A2) is reversed. In our case of expected yield $E[y \mid x]$ in (9), and common distribution up to location shift, when $\mu_t = z_t, t \in \Omega_N$, then (A2) becomes

(A3) $\quad (\mu_i - \mu_j)\int_0^{y^u}\left(\frac{f(y-\mu_i \mid x)}{\bar{F}(y-\mu_i \mid x)} - \frac{f(y-\mu_j \mid x)}{\bar{F}(y-\mu_j \mid x)}\right)\prod_{k\in\Omega_N}\bar{F}(y-\mu_k \mid x)dy \le 0.$

Statement (A3) is true whenever

(A4) $\quad (\mu_i - \mu_j)\left(\frac{f(y-\mu_i \mid x)}{\bar{F}(y-\mu_i \mid x)} - \frac{f(y-\mu_j \mid x)}{\bar{F}(y-\mu_j \mid x)}\right) \le 0 \quad \forall y \in [0, y^u].$

Under IFR then $\mathrm{Ln}[\bar{F}(y-\mu \mid x)]$ is concave in y for any relevant $\mu \in \mathbb{R}$. So IFR asserts that $f(y-\mu \mid x)/\bar{F}(y-\mu \mid x)$ is increasing in y and decreasing in location parameter μ. If $\mu_i \le (\ge)$ μ_j then

(A5) $\quad \frac{f(y-\mu_i \mid x)}{\bar{F}(y-\mu_i \mid x)} \ge (\le) \frac{f(y-\mu_j \mid x)}{\bar{F}(y-\mu_j \mid x)} \le 0 \quad \forall y \in [0, y^u],$

and (A4) follows. The assertion for DFR follows from the Schur-convexity version of inequality (A2).

Table 1. Yield Moments When Resource Availabilities Are Jointly Normally Distributed

$\rho \rightarrow$	Yield moments $\downarrow$	–0.5			0			0.5		
$(\mu_1, \mu_2) \rightarrow$		(10,10)	(8,10)	(12,10)	(10,10)	(8,10)	(12,10)	(10,10)	(8,10)	(12,10)
$(\sigma_1, \sigma_2) = (1,1)$	Mean	9.305	7.894	9.891	***9.435***	7.950	9.949	9.600	7.990	9.991
	Std. Dev.	0.721	0.885	0.880	***0.827***	0.959	0.938	0.912	1.000	0.988
	Skewness	-0.341	-0.335	-0.341	***-0.131***	-0.122	-0.144	-0.028	-0.022	-0.022
$(\sigma_1, \sigma_2) = (1,2)$	Mean	8.933	7.661	9.652	9.108	7.780	9.782	9.300	7.894	9.895
	Std. Dev.	1.179	0.934	1.592	1.303	1.020	1.696	1.420	1.077	1.835
	Skewness	-0.923	-0.463	-0.658	-0.664	-0.246	-0.448	-0.561	-0.229	-0.272
$(\sigma_1, \sigma_2) = (2,2)$	Mean	8.605	7.379	9.398	8.864	7.593	9.609	9.201	7.826	9.830
	Std. Dev.	1.452	1.555	1.537	1.642	1.751	1.744	1.829	1.908	1.915
	Skewness	-0.365	-0.379	-0.416	-0.177	-0.191	-0.176	-0.042	-0.049	-0.056
$(\sigma_1, \sigma_2) = (1,0.5)$	Mean	9.470	7.963	9.963	9.550	7.983	9.985	9.653	7.997	9.997
	Std. Dev.	0.584	0.936	0.473	0.654	0.963	0.492	0.707	0.999	0.503
	Skewness	-0.935	-0.275	-0.148	-0.671	-0.149	-0.029	-0.623	-0.052	-0.006
$(\sigma_1, \sigma_2) = (0.5,0.5)$	Mean	9.685	7.997	9.996	9.719	7.999	10.000	9.800	8.000	10.000
	Std. Dev.	0.361	0.488	0.494	0.416	0.500	0.500	0.459	0.499	0.504
	Skewness	-0.375	-0.064	-0.070	-0.144	-0.010	-0.011	-0.041	0.000	0.000