Journal of Agricultural and Resource Economics 34(1):34–52

Crop Yield Skewness and the Normal Distribution

David A. Hennessy

Empirical studies point to negative crop yield skewness, but the literature provides few clear insights as to why. This paper formalizes three points on the matter. Statistical laws on aggregates do not imply a normal distribution. Whenever the weather-conditioned mean yield has diminishing marginal product with respect to a weather-conditioning index, then there is a disposition toward negative yield skewness. This is because high marginal product in bad weather stretches out the yield distribution's left tail relative to that for weather. For disaggregated yields, unconditional skewness is decomposed into weather-conditioned skewness plus two other terms and each is studied in turn.

Key words: conditional distribution, crop insurance, negative skewness, normal distribution, statistical laws

Introduction

Crop yield distributions are used to model risk exposures, and also as an input in informed rating-making when designing and marketing crop insurance. As a review of recent literature on crop yield modeling should confirm, it is a controversial topic (Just and Weninger, 1999; Ker and Goodwin, 2000; Atwood, Shaik, and Watts, 2002, 2003; Sherrick et al., 2004). These papers disagree on whether crop yields should express nonzero skewness. At both conceptual and practical levels, the underlying problem arises from heterogeneity in growing conditions so that observed data are inevitably aggregated to some degree. Plots differ across space due to climate and soil variation, while different technologies may also be used. Concerns about the effects of spatial aggregation thus naturally arise.

A literature has emerged on problems with and methods to control for spatial aggregation difficulties when setting insurance rates (Wang and Zhang, 2002; Popp, Rudstrom, and Manning, 2005). There may also be confounding temporal and spatial heterogeneities in scrutinized data due to random weather events. Complete control over these heterogeneities is practically impossible, but clear theoretical underpinnings would certainly be helpful. Expanding our understanding of how yield aggregation affects empirical distributions is also important because the growing availability of weather derivatives may allow for effective yield risk management on the part of crop insurers at the aggregate level, but not on the part of growers (Woodard and Garcia, 2008).

Cases where skewness is likely positive were identified by Day (1965), Ramírez, Misra, and Field (2003), and also Chen and Miranda (2008), all for cotton. However, the

David A. Hennessy is professor, Department of Economics and Center for Agricultural and Rural Development, Iowa State University, Ames.

Review coordinated by David K. Lambert.

majority of studies have estimated negative skewness (Nelson and Preckel, 1989; Swinton and King, 1991; Moss and Shonkwiler, 1993; Ramírez, Misra, and Field, 2003; Atwood, Shaik, and Watts, 2002, 2003). This is at variance with the suggestion that the crop yield distribution should be normal, and so have zero skewness, in light of statistical limit considerations. Referring to a version of the central limit theorem (CLT), Just and Weninger (1999, p. 302) state, "Central limit theory thus calls into question the nonnormal empirical results found to date with aggregate time-series data." On the other hand, Goodwin and Mahul (2004, pp. 13–14) and others hold that spatial dependence and systemic risk factors indicate a straightforward application of central limit theory is not appropriate.[1] We see a need for clarification.

Apart from Just and Weninger (1999), two other papers have sought to provide foundations for the origin of observed yield skewness. Ker and Goodwin (2000, p. 465 and figure 1) reason that aggregating over some weather-conditioned yield distribution should predispose unconditional yield toward negative skewness whenever the weather variable is itself negatively skewed. Focusing on the small plot level of analysis, Hennessy (2009) relates yield skewness to the von Liebig law of the minimum technology, as studied in, e.g., Berck and Helfand (1990). When resource availabilities are stochastic, this technology can support positive or negative skewness. Efforts to tightly control the left tail of a resource availability can create a negative skewness tendency in yields. Irrigation, however, may increase skewness by completely eliminating a source of yield shortfall.

Assuming a general technology and dealing with aggregation issues, this paper makes three main points. We invoke the systemic-idiosyncratic risk decomposition concept that is widely used in financial risk theory (LeRoy and Werner, 2001).[2] The first point is that statistical large number and limit laws assert more about the distribution of an average than just convergence to normality. The relevance of statistical large number and limit laws to crop yield distributions has long been recognized (Hirshleifer, 1961). Still, the state of the crop yield distribution modeling literature suggests a need for clarification as to what these statistical laws actually imply for crop yields. In specifications relevant to modeling mean crop yields, the laws also provide a limiting value of zero for the distribution variance. In short, in the aggregate, only the distribution of systemic risk sources remains as a determinant of the crop yield distribution. We formally explain why CLT is relevant, but yet does not help in any way to narrow down the set of distributions appropriate for crop yield modeling.

The paper's second main point is constructive. It concerns the role of weather conditioning, and so can be seen as a formalization of and development on the graphical argument in Ker and Goodwin (2000). We clarify that, due to a result reported in van Zwet (1964), a beneficial random systematic weather factor with decreasing marginal product ensures the yield distribution will have smaller (i.e., more negative) skewness than that of the underlying weather factor. Intuitively this is because, in the change of random variables from weather factor to yield, the production function's concavity

[1] Wang and Zhang (2003) use dependent versions of the CLT, but for the purpose of estimating a crop insurance company's portfolio risk as the number of risks grows—not for identifying yield distributions. As we will explain, it is more meaningful to use CLT when studying the distribution of a sum than when studying the distribution of an average.

[2] Portfolio theory could, and sometimes does, seek to model spatial dimensions to systemic risk. In portfolio theory, of course, it is not geographic distance but rather attribute distance that is modeled, as when food company shares are grouped to be more similar to one another than to defense company shares.

stretches out the left tail of the weather factor distribution. This negative bent is consistent with the empirical literature on yield skewness.

Our third main point is to decompose and study skewness so as to better understand its origin when the level of analysis is sufficiently small that random effects do not average out. We relate skewness in aggregate yields to skewness in local yields, where local idiosyncrasies play a role. As in Hennessy (2009), the aggregate-local yield skewness relation is decomposed into three terms. Relative to the earlier paper, our contributions in the present paper are (*a*) to use main points 1 and 2 in the present paper to better understand the skewness decomposition, and (*b*) to sign the terms under various assumptions regarding the weather-yield relationship across different idiosyncratic states.

The paper's organization is as follows. We first provide an analysis under the assumption that production plots are homogeneous. One deduction is then probed for implications on yield skewness. Next, the case of plot heterogeneity is addressed. Upon introducing a disaggregated idiosyncratic risk component, an expression relating conditional skewness with unconditional skewness is presented and analyzed. A brief discussion concludes.

Analysis Under Yield Aggregation

For any plot of land, we consider three contributions to the yield realization:

- *Contribution I.* The land could differ innately from other plots due to endowments. This contribution is not random, and we will refer to it as spatial heterogeneity or land heterogeneity.
- *Contribution II.* There could be a systemic factor, such as weather, which is common to all plots, but realizations could differ from year to year. Then there would be a common yield across plots each year, but this yield would vary from year to year.
- *Contribution III.* There could be a local idiosyncratic, or LI, factor. This contribution involves different yield realizations for otherwise identical plots in the same year, perhaps because of local weather, pest, or infrastructure problems.

In this section it will be assumed that the yield observation is at an aggregated level, perhaps at the township or county level. We first jointly consider systemic (contribution II) and idiosyncratic sources of temporal randomness (contribution III). We then allow for contribution I, spatial heterogeneities.

Homogeneous Land

Our statistical model of crop yields is as follows. There are n land units of equal area in a defined geographic region. The land units are homogeneous in that they are agronomically identical, face the same systemic weather events, and are farmed in the same way. Yield is random in two senses. There is systemic randomness, but there are also local idiosyncrasies. First, conditional on some "weather favorability" index $w \in W = [0,1]$, yield on the ith land unit is the random realization $y_i(w)$ with support on some interval Y. This is contribution III.

These conditional draws $y_i(w)$, $i \in \{1, \ldots, n\}$, are independent and from the identical conditional distribution $F^{|w}(y_i): Y \rightarrow [0,1]$ with mean $\mu(w)$. Here the weather favorability index is given content by the assumption that $\mu(w)$ is a strictly increasing function. Define the inverse function as $w = \mu^{-1}(y)$. Second, contribution II is modeled by letting nature draw a w from distribution $G(w)$, where bounded density $g(w)$ exists everywhere on $w \in W$. The following ex ante, or weather unknown, yield distribution is identified in the appendix.

- PROPOSITION 1. *Suppose (a) land is homogeneous with yields given by independently distributed and identically drawn weather-conditioned draws from $F^{|w}(y_i)$, and (b) the weather-conditioned mean yield $\mu(w)$ exists and is strictly increasing. Let weather be distributed according to $G(w)$ on $w \in W$, where the density exists everywhere on the support. Then the region's limiting ex ante distribution, as $n \rightarrow \infty$, of mean crop yield is represented by $T(y)$, where $T(y) = G[\mu^{-1}(y)]$.*

Proposition 1 shows that LI are averaged out under the strong law of large numbers, leaving only the systemic component. The point here is that nonsystemic risks will be drowned out in the face of area-wide systemic weather risks so that the distribution of the systemic component dominates and there is no special role for the normal distribution. To clarify where confusion may lie on what proposition 1 conveys, let us first present where the normal distribution does arise in yield determination. To facilitate exposition, fix the value of w and suppress it in equations (1)–(3) to follow.

Suppose that x_i are independent draws from a common distribution with mean μ and standard deviation σ, while $\bar{x}_n \equiv (x_1 + \ldots + x_n)/n$ is the mean of n such draws. Then, and leaving technical definitions to the appendix, the central limit theorem asserts that the probability of $(\bar{x}_n - \mu)n^{0.5}/\sigma$ being no larger than some value x converges in distribution to the standard normal with mean 0 and variance 1; or

$$(1) \qquad P\left(\frac{(\bar{x}_n - \mu)n^{0.5}}{\sigma} \leq x\right) \overset{dist}{\rightarrow} \mathrm{Nor}_{(0,1)}(x).$$

However $(\bar{x}_n - \mu)n^{0.5}/\sigma$ is not in any way a measure of average yield, as represented by statistic $\bar{x}_n$. On the other hand, an implication of the strong law of large numbers is that the distribution of $\bar{x}_n$ converges as follows:

$$(2) \qquad P\left(\operatorname*{Lim}_{n \rightarrow \infty}(\bar{x}_n - \mu) \leq x\right) = \begin{cases} 0, & x < 0, \\ 1, & x \geq 0. \end{cases}$$

In order to reconcile (1) with (2), divide $(\bar{x}_n - \mu)n^{0.5}/\sigma$ in (1) by $n^{0.5}$ so that variance of the rescaled random variable becomes, at the limit, n^{-1} instead of 1. So (1) becomes

$$(3) \qquad P\left(\frac{(\bar{x}_n - \mu)}{\sigma} \leq x\right) \overset{dist}{\rightarrow} \mathrm{Nor}_{(0,n^{-1})}(x).$$

It follows from (3) that $\mathrm{Lim}_{n \rightarrow \infty} \mathrm{Nor}_{(0,n^{-1})}(x)$ has value 0 on $x < 0$, and value 1 on $x \geq 0$, which is consistent with (2). When all appropriate conditioning factors have been controlled for, mean yield converges to a degenerate distribution as described in (2). By use of the stabilizing scaling function $n^{0.5}/\sigma$, (1) provides a description of how convergence occurs. Only when multiplied by the artificially introduced scaling factor $n^{0.5}/\sigma$ will mean yield follow the standard normal distribution. Finally, when index w is allowed to follow a nondegenerate distribution, then expression (2) above ensures that the distribution of w determines the distribution of yield in the manner given by proposition 1, $T(y) = G[\mu^{-1}(y)]$.

One may wonder how robust this conclusion is when crop yields are clearly positively correlated over space. Note first when thinking this through that much if not most of this dependence may not persist after conditioning on a systemic weather factor. Suppose for the moment one can control for a systemic factor and that there only remain dependent LI. Central limit theory results allowing for spatial dependence among LI exist. Indeed, Wang and Zhang (2003), in their study of crop insurance risk pooling opportunities, have allowed for finite range dependence whereby yields are assumed to be independent beyond a defined spatial distance. Under finite range dependence, the variance of mean yield still converges to zero. In their empirical study, Wang and Zhang used county data for wheat, soybeans, and corn across 2,000 to 2,640 counties in the United States over 1972–1997. They found the distance for any positive dependence to be at most 570 miles, and in many cases much less. They did not control for weather, so their distribution is largely due to weather and not to weather-conditioned idiosyncratic risks. Their range conclusions are similar to those found by Goodwin (2001) for Illinois, Indiana, and Iowa corn production, where spatial correlation declines to 0.1 by 200 miles in typical years and by 400 miles in drought years.

Of what relevance is all this when interpreting proposition 1? It depends on how data are being aggregated. If yield data are from a small area, then one can reasonably assume a common w realization for all land at issue. If data are from a larger area, then a spatial stochastic process is appropriate. Let $w(i,j)$ and $y(i,j)$ be weather and yield random variables at map coordinate (i, j), $(i, j) \in \{1, \ldots, I\} \times \{1, \ldots, J\}$, where weather follows the arbitrary joint distribution $G[w(1,1), w(1,2), \ldots, w(I, J-1), w(I, J)]$. Then $w(i,j) = \mu^{-1}[y(i,j)]$, and yield follows the joint distribution $G[\mu^{-1}[y(1,1)], \mu^{-1}[y(1,2)], \ldots, \mu^{-1}[y(I, J-1)], \mu^{-1}[y(I, J)]]$. Introducing the spatial dependencies adds a layer of notation to the finding in proposition 1, but does not change the underlying result.

Local weather variations are a large component of local idiosyncrasies. It is difficult to disentangle the two when presenting a formal model of spatio-temporal yield randomness, and even more so as a practical concern. But no matter how one looks at it, the normal distribution should have no special role in understanding how averages are distributed. Suppose (*a*) all practical conditioning problems have been sorted out, including the thorny issue of spatial correlations in weather, and (*b*) a central limit theorem applies to what is left. Then mean yield will have zero variance at the limit. More positively, while we may not know how to condition well or have the requisite data, the observed distribution should be seen to reflect what has not been conditioned on and not to have been formed by central limit theory.

Upon reflection, one may conclude that proposition 1 is an almost trivial (if formal) application of statistical laws. We do not disagree, but see its merits on two fronts. First, yield random variables are realizations of an unknown spatio-temporal process where yields at each point in space need further conditioning to account for relevant agronomic attributes. Given the complexity of the context, exactly what to condition on can be lost in the mix; consequently, how statistical laws apply warrants formal delineation. In addition, the procedure of formal delineation reveals a relationship that is important in understanding the yield distribution, namely $T(y) = G[\mu^{-1}(y)]$.

Notice the process of arriving at $T(y)$ in proposition 1 is an example of that ubiquitous statistical technique, the change of variables. Our paper's second main point arises from a consideration of what effect this change of variables has on the skewness, as the weather variable affects the yield variable. For some random variable η, and with $\mathrm{E}[\cdot]$ as the expectation operator, recall that the skewness statistic is:

$$\Upsilon(\eta) = \frac{\mathrm{E}\left[(\eta - \mathrm{E}[\eta])^3\right]}{\left\{\mathrm{E}\left[(\eta - \mathrm{E}[\eta])^2\right]\right\}^{3/2}}. \tag{4}$$

In theorem 2.2.1 on p. 10, and later remarks on p. 16 of his seminal work on convex transformations of random variables, van Zwet (1964) established the following. Provided the skewness statistics exist, then any transformed random variable $\xi = L(\eta)$ has a smaller (larger) skewness statistic than η whenever $L(\eta)$ is increasing and concave (convex). Thus we have:

- PROPOSITION 2. *Make the assumptions of proposition 1. Assume too that $\mu(w)$ is concave (convex) and the skewness statistics exist. Then yield skewness is smaller (larger) than weather index skewness, or*

$$\Upsilon[\mu(w)] = \frac{\mathrm{E}\left[(\mu(w) - \mathrm{E}[\mu(w)])^3\right]}{\left\{\mathrm{E}\left[(\mu(w) - \mathrm{E}[\mu(w)])^2\right]\right\}^{3/2}} \le (\ge) \frac{\mathrm{E}\left[(w - \mathrm{E}[w])^3\right]}{\left\{\mathrm{E}\left[(w - \mathrm{E}[w])^2\right]\right\}^{3/2}} = \Upsilon(w). \tag{5}$$

Proposition 2 states that when mean yield is an increasing and concave function of, say, growing degree days, then mean crop yield is more negatively skewed than is growing degree days. In particular, when growing degree days has zero skewness, then mean crop yield will have negative skewness. To develop intuition on this, consider the probability weights given to weather index intervals $[w^l, w^l + \delta]$ and $[w^h, w^h + \delta]$, where $w^h > w^l$ and $\delta > 0$. For $j \in \{l, h\}$, the probability weight on interval $[w^j, w^j + \delta]$ is transformed to apply over yield interval $[\mu(w^j), \mu(w^j + \delta)]$, where concavity ensures that $\mu(w^l + \delta) - \mu(w^l) > \mu(w^h + \delta) - \mu(w^h)$. The transformation stretches the left tail density weightings over a comparatively larger interval in the yield variable and contracts the right tail weightings. This creates a tendency toward the sort of left-tail to right-tail asymmetry associated with negative skewness. Figure 1 provides an illustration.

An alternative means of making the same observation is through differentiating $T(y)$ to obtain density:

$$t(y) = g\left[\mu^{-1}(y)\right]\frac{d\mu^{-1}(y)}{dy} = \left.\frac{g(w)}{\partial\mu(w)/\partial w}\right|_{w=\mu^{-1}(y)}. \tag{6}$$

As the denominator $\partial\mu(w)/\partial w$ is declining in w, the yield density at the same quantile, or upon using the change of variables $w = \mu^{-1}(y)$, is increasing in w. In other words, the yield distribution's left tail density is stretched longer and thinner upon mapping from weather domain to yield domain.

Note a concave $\mu(w)$ would suggest a cardinal interpretation of the weather favorability index in that expected yield has decreasing marginal product in the index. The weather favorability indices that come to mind, such as growing degree days or the moisture stress index, are cardinal representations of scientific measures. In light of

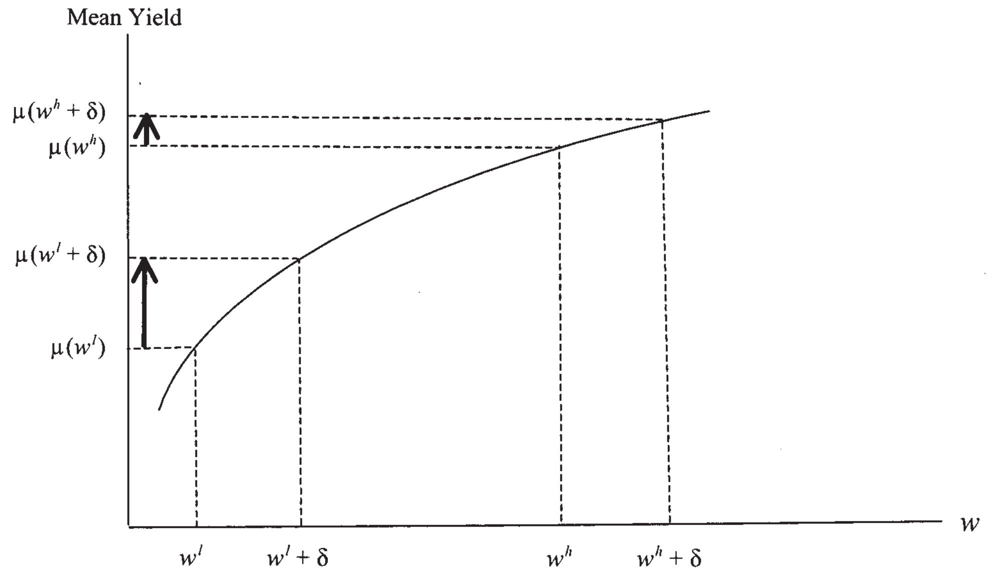

Figure 1. Mapping weather variation to yield variation

findings in Schlenker (2006), Schlenker and Roberts (2006), and Schlenker, Hanemann, and Fisher (2006), it may be necessary to include a separate index for harmfully high degree days.

A second but related issue is whether concavity and monotonicity properties apply. Because weather extremes are seldom good for crops suited to a given climatic region, yield will not be monotone in precipitation or temperature (Thompson, 1986; Tannura, Irwin, and Good, 2008). It should, however, be monotone increasing in a well-designed weather favorability index. There remains the issue of concavity. Although using disparate approaches and contexts, and although not seeking evidence on concavity, Riha, Wilks, and Simoens (1996), Porter and Semenov (1999, 2005), Schlenker (2006), Almaraz et al. (2008), and others find that an increase in within-year weather variance reduces average crop yield. This suggests an overall concave shape.

Before proceeding, and in particular reference to the two preceding paragraphs, a comment is in order regarding relation (6). In statistical terms, a change of variables between the probability distribution on weather and that on yield has been conducted (Bain and Engelhardt, 1992, pp. 197–209). The change of variable technique is most directly applicable when the underlying map, in this case from weather to yield, is a univariate and strictly monotone function. In our case, the underlying map is not generally likely to be either univariate or monotone. For the case of corn, both temperature and precipitation for each month over the growing season have been shown to be relevant (Thompson, 1986; Tannura, Irwin, and Good, 2008). There also are likely to be nonmonotonicities in yield responses to temperature (Schlenker and Roberts, 2006).

However, the change of variable method generalizes to allow for multivariate and nonuniform relationships so long as sufficient information is available on the relationship (Bain and Engelhardt, 1992, pp. 202–209). Stated differently, more general versions of proposition 1 are available where the same overall insight would continue to apply. Empirical enquiry in the area is likely to be challenging. To illustrate just one problem, to the extent that corn breeding programs have emphasized drought tolerance over the past 20 years, the relationship will be time dependent (Staggenborg, Dhuyvetter, and Gordon, 2008; Fatka, 2009).

As to the distribution of the underlying weather index, the matter is important for a wide variety of industries, including power and outdoor entertainment sectors. These are among the main users of weather derivative markets, and participants in these markets have a strong pecuniary interest in identifying appropriate distribution forms for valuation of derivatives on such indices as cooling degree days. The skewness statistic for any weather variable is as likely to depend on location as are mean and variance. The weather derivative valuation literature has not yet developed to the point where identifying and accounting for stylized features on skewness have occurred; see, e.g., Jewson and Brix (2005) where modeling skewness does not arise. This is in contrast with derivatives on stock market prices and, of course, with rate setting in crop insurance markets.

To conclude our study of proposition 2, we provide and discuss a weather-yield relationship that supports a closed-form solution for yield skewness when the weather index is normally distributed. The example raises the possibility that there is more to the skewness transformation than proposition 2 asserts.

Example 1. Consider a single-input Mitscherlich-Baule production relation between mean yield and weather:

$$\mu(w) = \lambda_0 - \lambda_1 e^{-\lambda_2 w}, \tag{7}$$

where $w > 0$, $\lambda_0 \geq \lambda_1 > 0$, $\lambda_2 > 0$ to ensure a positive, increasing, concave relation. Let w be normally distributed with mean θ and variance σ^2 so that skewness for the weather distribution is zero.[3] As for yield skewness, use of moment-generating functions for the normal distribution shows:

$$\Upsilon\left[\mu(w)\right] = -\left(e^{\lambda_2^2\sigma^2} - 1\right)^{0.5}\left(e^{\lambda_2^2\sigma^2} + 2\right) < 0, \tag{8}$$

where details are provided in the appendix. The function is decreasing in both the relative curvature parameter λ_2 and the weather index variance parameter σ^2. That $\Upsilon[\mu(w)]$ is a decreasing function of λ_2 in this model suggests, in addition to the concavity status of $\mu(w)$, the relative curvature of $\mu(w)$ may be important in determining the extent to which yield skewness differs from weather index skewness.

Heterogeneous Land

In this subsection, the model is extended so that land is no longer homogeneous. In particular, land attribute $z \in Z = [0, 1]$ follows a continuously differentiable distribution $M(z)$. Conditional on any given value of z, there are n land units of equal area in a defined geographic region. Yield on the ith land unit is $y_i(w; z)$ These conditional draws are independent and from the identical conditional distribution $F^{|w,z}(y)$ with mean $\mu(w, z)$,

---

[3] As is typical when working with the normal distribution, the left tail problem of w [such that $\mu(w) < 0$] is assumed to be negligible and is ignored.

which is assumed to be increasing and differentiable in both arguments. Define the attribute-conditioned inverse as $w = h(y; z)$ and the weather-conditioned inverse as $z = r(y; w)$.

■ PROPOSITION 3. *Let (a) land be heterogeneous with yields given by i.i.d. weather- and yield-conditioned draws from $F^{|w,z}(y)$; (b) the weather and land attribute conditional mean yield $\mu(w, z)$ exists. Let weather and land attributes be distributed according to $G(w)$: $W \to [0,1]$ and $M(z)$: $Z \to [0,1]$, respectively, in each case with densities defined over the entire support. Then, as $n \to \infty$, the region's*

(i) *limiting mean crop yield distribution, ex post for given w, is represented by $L^{|w}(y)$, where $L^{|w}(y) = M[r(y; w)]$;*

(ii) *limiting mean crop yield distribution for a given land attribute value z is represented by $T^{|z}(y)$, where $T^{|z}(y) = G[h(y; z)]$; and*

(iii) *unconditional limiting mean crop yield distribution is represented by $T(y)$, where*

$$T(y) = \int_Z G\big[h(y;z)\big]dF(z) = \int_W M\big[r(y;w)\big]dG(w).$$

Again, in each case, the limiting distribution is that where the idiosyncratic randomness has been ignored. The usual left-tail problem aside, the normal distribution may or may not result. But it has no special place in the analysis.

Relating Conditional and Unconditional Moments

Sometimes the level of analysis will be such that local idiosyncrasies do not average out. One may then wonder what additional facets of the production environment become relevant and where the result in proposition 1 fits in at the more disaggregated level. In order to bring further clarity to the role of data aggregation in determining yield skewness, this section introduces a local idiosyncratic (LI) factor that is not averaged away, so the scale of analysis is at the small plot level. While working primarily at the disaggregated level, Hennessy (2009) identified a relationship between skewness at the aggregate and disaggregate levels. We will provide an alternative approach to establishing the relationship, show where our earlier propositions fit into the relationship, and offer some comments about how conditional production technologies may affect the relationship.

The marginal distribution for yield at some arbitrarily chosen ith location with fixed z attributes is $F(y)$: $Y \to [0,1]$, where the location and z indicators are henceforth dropped in order to avoid notational clutter. The unconditional mean yield is

$$\mu = \int_W \mu(w)\,dG(w),$$

and the conditioned deviation in mean yield is $\delta(w) = \mu(w) - \mu$. The residual between ith location yield and the conditional mean is $\varepsilon(w) = y - \mu(w)$. This difference arises from LI. Of course,

$$\int_Y \varepsilon(w)\,dF^{|w}(y) = 0.$$

We are interested in comparing the conditional and unconditional expectations of some mean-normalized function,

$$(9) \qquad S(y-\mu) = S\big[\varepsilon(w) + \delta(w)\big],$$

where the functions of interest will all be of kth moment form $[\varepsilon(w) + \delta(w)]^k$. With $\partial^j S(\cdot)/\partial y^j$ as the jth derivative of function $S(\cdot)$, a third-order Taylor's series expansion around $\varepsilon(w) = 0$ identifies

$$(10) \qquad S\big[\varepsilon(w) + \delta(w)\big] \approx S(\cdot)\big|_{\varepsilon(w)=0} + \varepsilon(w)\frac{\partial S(\cdot)}{\partial y}\bigg|_{\varepsilon(w)=0} + \frac{\big[\varepsilon(w)\big]^2}{2}\frac{\partial^2 S(\cdot)}{\partial y^2}\bigg|_{\varepsilon(w)=0} + \frac{\big[\varepsilon(w)\big]^3}{6}\frac{\partial^3 S(\cdot)}{\partial y^3}\bigg|_{\varepsilon(w)=0}.$$

Use the notation for kth conditional and unconditional moments:

$$(11) \qquad \text{Conditional:} \quad C^{(k)}(y\,|\,w) = \int_Y \big[y-\mu(w)\big]^k dF^{|w}(y),$$

$$\text{Unconditional:} \quad C^{(k)}(y) = \int_W \int_Y \big[y-\mu\big]^k dF^{|w}(y)\, dG(w) = \int_Y \big[y-\mu\big]^k dF(y).$$

Noting that $C^{(1)}(y \mid w) = 0$, insert (10) as $[y-\mu]^k$ into the $C^{(k)}(y)$ equation in (11) to establish:

$$(12) \qquad C^{(k)}(y) \approx \int_W S\big[\delta(w)\big]dG(w) + \frac{1}{2}\int_W \frac{\partial^2 S(\cdot)}{\partial y^2}\bigg|_{\varepsilon(w)=0} C^{(2)}(y\,|\,w)\,dG(w) + \frac{1}{6}\int_W \frac{\partial^3 S(\cdot)}{\partial y^3}\bigg|_{\varepsilon(w)=0} C^{(3)}(y\,|\,w)\,dG(w).$$

Now let[4]

$$(13) \qquad S(y-\mu) = \big[y-\mu(w) + \delta(w)\big]^3,$$

so that the third-order approximation in (10) is exact. Relation (12) becomes

$$(14) \qquad C^{(3)}(y) = \int_W \big[\delta(w)\big]^3 dG(w) + 3\int_W \delta(w) C^{(2)}(y\,|\,w)\,dG(w) + \int_W C^{(3)}(y\,|\,w)\,dG(w),$$

as identified in Hennessy (2009). Relative to that paper, our contribution is to sign the terms in expression (14). In particular, we consider two states of the world, good and bad, in addition to weather index w.

Suppose, conditional on w, there can be only two states of nature where w is assumed to be some beneficial growth factor such as growing degree days. There can be good overall growing conditions in the locality with probability π and yield $y^g(w)$, and also bad overall growing conditions with probability $1-\pi$ and yield $y^b(w)$. Here g could represent

[4] Conditional and unconditional kurtosis and other statistics can be compared in similar manner.

good planting conditions, perhaps because some fields among a set of identical fields were planted just before a rain event that delayed planting on the remainder for a week. Of course, $y^g(w) \geq y^b(w)\ \forall\ w \in W$, while in addition, $\partial y^j(w)/\partial w \geq 0\ \forall\ w \in W, j \in \{b, g\}$. Some algebra confirms:

$$\begin{aligned}
(15)\qquad & \mu(w) = \pi y^g(w) + (1-\pi)y^b(w),\\
& C^{(2)}(y \mid w) = \pi(1-\pi)\big[y^g(w) - y^b(w)\big]^2,\\
& C^{(3)}(y \mid w) = \pi(1-\pi)(1-2\pi)\big[y^g(w) - y^b(w)\big]^3.
\end{aligned}$$

Further observations about equation (14) are that

$$\int_W \delta(w)\, dG(w) = 0,$$

meaning that the well-known covariance decomposition leads to

$$\begin{aligned}
(16)\qquad & \int_W \delta(w) C^{(2)}(y \mid w)\, dG(w)\\
&= \int_W \delta(w)\, dG(w) \int_W C^{(2)}(y \mid w)\, dG(w) + \mathrm{Cov}\big[\delta(w), C^{(2)}(y \mid w)\big]\\
&= \mathrm{Cov}\big[\mu(w) - \mu, C^{(2)}(y \mid w)\big]\\
&= \mathrm{Cov}\big[\mu(w), C^{(2)}(y \mid w)\big],
\end{aligned}$$

since constant μ has zero covariance with any other variable.

In light of the points made above, equation (14) may be rewritten as:

$$\begin{aligned}
(17)\qquad C^{(3)}(y) = & \overbrace{\int_W \big[\delta(w)\big]^3 dG(w)}^{\text{Term I}}\\
& + \overbrace{3\pi(1-\pi)\mathrm{Cov}\Big[\pi y^g(w) + (1-\pi)y^b(w), \big[y^g(w) - y^b(w)\big]^2\Big]}^{\text{Term II}}\\
& + \overbrace{\pi(1-\pi)(1-2\pi)\int_W \big[y^g(w) - y^b(w)\big]^3 dG(w)}^{\text{Term III}}.
\end{aligned}$$

Each term is examined in turn below.

Term I. This expression has already been studied in proposition 2 and the discussions surrounding it. It is the systemic component of the unconditional third central moment. It would be the third central moment were the level of analysis sufficiently aggregated.

Term II. By the covariance inequality, term II is negative whenever one expression is increasing in w while the other is decreasing in w. Now $\partial\mu(w)/\partial w \geq 0$ as $\partial y^j(w)/\partial w \geq 0\ \forall\ w \in W, \forall j \in \{b, g\}$, and $\pi \in [0, 1]$. The second term in the covariance expression has derivative $2[y^g(w) - y^b(w)]\{\partial y^g(w)/\partial w - \partial y^b(w)/\partial w\}$. This is negative if the LI and systemic risk factors substitute, but positive if they complement. Our present knowledge on crop production theory suggests both are possible where the crop in question, location, and production practices will factor into determining the sign. Some examples illustrate.

- *Case A.* If $y^g(w) = y^b(w) + \lambda\ \forall\ w \in W$, where $\lambda > 0$, then $\partial y^g(w)/\partial w \equiv \partial y^b(w)/\partial w$ and term II has value zero.

- *Case B.* If instead $y^g(w) = y^b(w + \lambda)\ \forall\ w \in W$, where $\lambda > 0$ and $y^b(\cdot)$ is strictly concave, then $\partial y^g(w)/\partial w < \partial y^b(w)/\partial w\ \forall\ w \in W$ and term II is negative.

- *Case C.* Finally consider a special instance of the von Liebig technology as specified in Berck and Helfand (1990). Let $y(w, \varepsilon) = \min[w, \varepsilon]$, where $\varepsilon^g > \varepsilon^b$, $\varepsilon = \varepsilon^g$ with probability π, and $\varepsilon = \varepsilon^b$ with probability $1 - \pi$. Then[5]

$$(18)\qquad y^g(w) = \min[w, \varepsilon^g], \qquad y^b(w) = \min[w, \varepsilon^b],$$

$$\frac{\partial y^g(w)}{\partial w}\begin{cases} = 1 & \text{if } w < \varepsilon^g \\ \in [0,1] & \text{if } w = \varepsilon^g \\ = 0 & \text{if } w > \varepsilon^g \end{cases}, \qquad \frac{\partial y^b(w)}{\partial w}\begin{cases} = 1 & \text{if } w < \varepsilon^b \\ \in [0,1] & \text{if } w = \varepsilon^b \\ = 0 & \text{if } w > \varepsilon^b \end{cases}.$$

Thus,

$$(19)\qquad \frac{\partial y^g(w)}{\partial w} - \frac{\partial y^b(w)}{\partial w}\begin{cases} = 0 & \text{if } w \in [0, \varepsilon^b) \cup (\varepsilon^g, \infty) \\ \in [0,1] & \text{if } w \in \{\varepsilon^b, \varepsilon^g\} \\ = 1 & \text{if } w \in (\varepsilon^b, \varepsilon^g) \end{cases}.$$

Clearly, complementarity between weather and maximum output ensures that $\partial y^g(w)/\partial w - \partial y^b(w)/\partial w$ is nowhere strictly negative on $w \in W$ so that term II is nonnegative.

Term III. If good growing conditions are more common, or $\pi > 0.5$, then the monotonicity condition $y^g(w) \geq y^b(w)\ \forall\ w \in W$ ensures that term III is negative. A comparatively heavy weighting on the good state disposes the LI component of yield toward negative skewness, while $\pi < 0.5$ ensures positive skewness. Stepping outside the realm of LI and into that of technical change, term III has a practical interpretation. As previously mentioned, corn breeding programs have placed emphasis on drought tolerance in recent years. This may be seen as an endeavor to increase the good growing conditions probability, π.

Figure 2 illustrates case B above. In the figure the production functions are the same except that $y^b(w)$ is a rightward translation of $y^g(w)$, parallel to the w axis. Due to concavity, the vertical gap between outputs under the two states declines as weather improves from any index value y^l to another value w^h such that $w^h > w^l$. Weather-conditioned mean yield increases with w, while conditional variance decreases with w. Somewhat similar to the analysis in proposition 2 and figure 1, introducing LI creates more dispersion on the left tail of the yield distribution, and this is what generates the tendency toward negative skewness. In contrast, for case C, the LI creates more dispersion on the right tail. In spreading out the right tail realizations, it promotes positive skewness.

[5] The $\in [0, 1]$ component merely provides bounds on the derivative at the kink point, where the derivative from above is 0 and the derivative from below is 1. Any line through the kink point with slope in $[0, 1]$ does not intersect the production function.

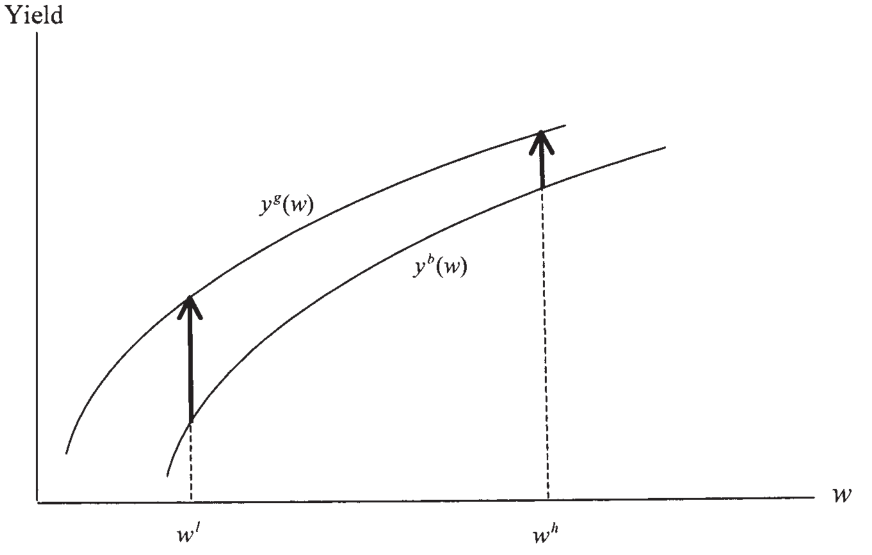

Figure 2. Yield under good and bad local idiosyncrasies

A final comment about decomposition (17) is that term II disappears whenever the conditional variance is independent of the conditioner. This is the case in figure 1 of Ker and Goodwin (2000), where a mixture of zero skew weather-conditioned yield distributions is represented. The conditioner itself is negatively skewed. This negative skewness (our term I) ensures negative unconditional yield skewness because Ker and Goodwin do not include an idiosyncratic component (our term III) in their graphical model.

Conclusion

With the ultimate goal of moving the literature toward developing and testing hypotheses on contexts in which positive or negative yield skewness should be observed, this paper has sought to do three things. First, it has pointed out that while both the strong law of large numbers and the central limit theorem should indeed apply given adequate approximation of the relevant statistical assumptions, there is a catch. Yield variance, when appropriately conditioned, should recede to zero and the limiting distribution is degenerate. This means that when systemic heterogeneities exist in the data under consideration, these will dominate to determine the shape of the yield distribution. Second, we identify an effect that tilts the skewness of aggregate yield to be more negative than weather factor skewness whenever the weather factor expresses positive but diminishing marginal impact on aggregated mean yield. Finally, moving to disaggregated yields generates two further effects. We show that both could act to increase or decrease yield skewness at the small plot level of analysis.

The paramount question when conditioning yield distributions is to what end the results will be applied. For the purpose of crop insurance, conditioners should be what can be observed at economically low cost by the insurer before contractual agreement. Ideally this will include clay content, water holding capacity, and other agronomic characteristics, as well as pre-season weather variables such as soil moisture or El Niño events. The growing literature on yield distributions seeking to account for spatial effects has yet to control for land heterogeneity beyond the use of location indicators (e.g., Ozaki et al., 2008). This is understandable because the focus is on other aspects of systemic variation. But improvements in geographical information systems and efforts by yield modelers in the climate change literature suggest that such conditioning should be feasible. Indeed, Gergaud and Ginsburgh (2008) have conditioned on vineyard geographical and soil attributes when seeking to identify their roles in wine quality determination. As our propositions show, in order to better explain what remains after careful conditioning, more attention must be paid to how weather variables shape aggregated yield distributions.

[*Received June 2008; final revision received February 2009.*]

References

Almaraz, J. J., F. Mabood, X. Zhou, E. G. Gregorich, and D. L. Smith. "Climate Change, Weather Variability, and Corn Yield at a Higher Latitude Locale: Southwestern Quebec?" *Climatic Change* 88,2(2008):187–197.

Atwood, J., S. Shaik, and M. Watts. "Can Normality of Yields Be Assumed for Crop Insurance?" *Can. J. Agr. Econ.* 50,2(2002):171–184.

———. "Are Crop Yields Normally Distributed? A Reexamination." *Amer. J. Agr. Econ.* 85,4(2003): 888–901.

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

Berck, P., and G. Helfand. "Reconciling the von Liebig and Differentiable Crop Production Functions." *Amer. J. Agr. Econ.* 72,4(1990):985–996.

Capiński, M., and E. Kopp. *Measure, Integral, and Probability*, 2nd ed. London: Springer Publications, 2004.

Chen, S.-L., and M. J. Miranda. "Modeling Texas Dryland Cotton Yields, with Application to Crop Insurance Actuarial Rating." *J. Agr. and Appl. Econ.* 40,1(2008):239–252.

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

Fatka, J. "Monsanto Pipeline 'Promising.'" *Feedstuffs* (12 January 2009), p. 5.

Gergaud, O., and V. Ginsburgh. "Natural Endowments, Production Technologies, and the Quality of Wines in Bordeaux. Does Terroir Matter?" *Economic J.* 118, special issue (2008):F142–F157.

Goodwin, B. K. "Problems with Market Insurance in Agriculture." *Amer. J. Agr. Econ.* 83,3(2001): 643–649.

Goodwin, B. K., and O. Mahul. "Risk Modeling Concepts Relating to the Design and Rating of Agricultural Insurance Contracts." Policy Res. Work. Paper No. 3392, World Bank, Washington, DC, September 2004.

Greene, W. H. *Econometric Analysis*, 5th ed. Upper Saddle River, NJ: Prentice-Hall, Inc., 2003.

Hennessy, D. A. "Crop Yield Skewness Under Law of the Minimum Technology." *Amer. J. Agr. Econ.* 91,1(2009):197–208.

Hirshleifer, J. "Risk, the Discount Rate, and Investment Decisions." *Amer. Econ. Rev.* 51,2(1961): 112–120.

Jewson, S., and A. Brix. *Weather Derivative Valuation: The Meteorological, Statistical, Financial, and Mathematical Foundations.* Cambridge, UK: Cambridge University Press, 2005.

Just, R. E., and Q. Weninger. "Are Crop Yields Normally Distributed?" *Amer. J. Agr. Econ.* 81,2(1999): 287–304.

Ker, A. P., and B. K. Goodwin. "Nonparametric Estimation of Crop Insurance Rates Revisited." *Amer. J. Agr. Econ.* 83,2(2000):463–478.

Lamperti, J. W. *Probability: A Survey of Mathematical Theory*, 2nd ed. New York: John Wiley and Sons, 1996.

LeRoy, S. J., and J. Werner. *Principles of Financial Economics.* Cambridge, UK: Cambridge University Press, 2001.

Lewin, J. W. "Some Applications of the Bounded Convergence Theorem for an Introductory Course in Analysis." *Amer. Mathematical Monthly* 94,10(1987):988–993.

Moss, C. B., and J. S. Shonkwiler. "Estimating Yield Distributions with a Stochastic Trend and Nonnormal Errors." *Amer. J. Agr. Econ.* 75,4(1993):1056–1062.

Nelson, C. H., and P. V. Preckel. "The Conditional Beta Distribution as a Stochastic Production Function." *Amer. J. Agr. Econ.* 71,2(1989):370–378.

Ozaki, V. A., S. K. Ghosh, B. K. Goodwin, and R. Shirota. "Spatio-Temporal Modeling of Agricultural Yield Data with an Application to Pricing Crop Insurance Contracts." *Amer. J. Agr. Econ.* 90,4(2008): 951–961.

Popp, M., M. Rudstrom, and P. Manning. "Spatial Yield Risk Across Region, Crop, and Aggregation Method." *Can. J. Agr. Econ.* 53,2-3(2005):103–115.

Porter, J. R., and M. A. Semenov. "Climate Variability and Crop Yields in Europe." *Nature* 400, no. 6746 (1999):724.

———. "Crop Response to Climatic Variation." *Philosophical Transactions of the Royal Society, B-Biological Sciences* 360, no. 1463 (2005):2021–2035.

Ramírez, O. A., S. Misra, and J. Field. "Crop Yield Distributions Revisited." *Amer. J. Agr. Econ.* 85,1(2003):108–120.

Riha, S. J., D. S. Wilks, and P. Simoens. "Impact of Temperature and Precipitation Variability on Crop Model Predictions." *Climatic Change* 32,3(1996):293–311.

Schlenker, W. "Inter-annual Weather Variation and Crop Yields." Unpub. work. paper, Dept. of Econ. and School of Internat. and Public Affairs, Columbia University, July 2006.

Schlenker, W., W. M. Hanemann, and A. C. Fisher. "The Impact of Global Warming on U.S. Agriculture: An Econometric Analysis of Optimal Growing Conditions." *Rev. Econ. and Statis.* 88,1(2006): 113–125.

Schlenker, W., and M. J. Roberts. "Nonlinear Effects of Weather on Corn Yields." *Rev. Agr. Econ.* 28,3(2006):391–398.

Sherrick, B. J., F. C. Zanini, G. D. Schnitkey, and S. H. Irwin. "Crop Insurance Valuation Under Alternative Yield Distributions." *Amer. J. Agr. Econ.* 71,2(2004):406–419.

Staggenborg, S. A., K. C. Dhuyvetter, and W. B. Gordon. "Grain Sorghum and Corn Comparisons: Yield, Economic, and Environmental Responses." *Agronomy J.* 100,6(2008):1600–1604.

Swinton, S. M., and R. P. King. "Evaluating Robust Regression Techniques for Detrending Crop Yield Data with Nonnormal Errors." *Amer. J. Agr. Econ.* 73,2(1991):446–451.

Tannura, M. A., S. H. Irwin, and D. L. Good. "Weather, Technology, and Corn and Soybean Yields in the U.S. Corn Belt." Marketing and Outlook Res. Rep. No. 2008-01, Dept. of Agr. and Consumer Econ., University of Illinois at Urbana-Champaign, February 2008.

Thompson, L. M. "Climatic Change, Weather Variability, and Corn Production." *Agronomy J.* 78,4(1986):649–653.

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

Wang, H. H., and H. Zhang. "Model-Based Clustering for Cross-Sectional Time-Series Data." *J. Agr., Biological, and Environ. Statis.* 7,1(2002):107–127.

———. "On the Possibility of a Private Crop Insurance Market: A Spatial Statistics Approach." *J. Risk and Insurance* 70,1(2003):111–124.

Woodard, J. D., and P. Garcia. "Weather Derivatives, Spatial Aggregation, and Systemic Risk: Implications for Reinsurance Hedging." *J. Agr. and Resour. Econ.* 33,1(2008):34–51.

Appendix:
Technical Definitions and Proof of Propositions

Relevant Large Number and Limit Laws

Let $P(\cdot)$ be a probability measure on some probability space $[\Omega, S, P(\cdot)]$, where $P(A)$ is the probability attached to condition set $A \subseteq S$.[6] Our analysis depends crucially upon statistical large number theory, where we choose to emphasize the standard strong law. Let $x_1, x_2, \ldots$ be independent random variables having the same probability distribution, i.e., they are independently drawn from identical distributions (i.i.d.).

Strong Law of Large Numbers (SLLN) (Lamperti, 1996, p. 50). For i.i.d. random variables x_1, $x_2, \ldots$ with distribution $F(x_i)$, assume their expected value, $\mu = \int x_i dF(x_i)$, exists. Then

(A1)
$$P\left(\lim_{n \to \infty} \frac{x_1 + \ldots + x_n}{n} = \mu \right) = 1.$$

Condition (A1) is also commonly referred to as almost sure convergence. Essentially, this condition asserts that the probability the sample average

$$n^{-1} \sum_{i=1}^{n} x_i$$

does not converge to true mean μ is negligible. The implications are very strong; the law allows for Monte Carlo methods to identify asymptotically exact estimates of the true expectation of some well-behaved function $f(x)$, i.e.,

$$\mathrm{E}\left[f(x)\right] = \lim_{n \to \infty} n^{-1} \sum_{i=1}^{n} f(x_i),$$

where the x_i are independent draws from a common distribution and $\mathrm{E}[\cdot]$ is the expectation operator with respect to that distribution.

While the law establishes what the mean converges to (almost surely), it provides very limited information on how convergence occurs. That is left to the central limit theorem, which describes the distribution a sample mean is likely to follow when the sample size is large. Distribution being a function, to characterize a central limit theorem one must have a concept of how to measure convergence between functions.

Convergence in Distribution (Bain and Engelhardt, 1992, p. 71). Probability measure $\upsilon_n(x)$ is said to converge in distribution to measure $\upsilon(x)$, written as

$$\upsilon_n \xrightarrow{dist} \upsilon,$$

if

(A2)
$$\lim_{n \to \infty} \upsilon_n(x) = \upsilon(x)$$

for all values on the domain of x at which $\upsilon(x)$ is continuous. The normal distribution is continuous everywhere, so the issue of continuity is moot.

Central Limit Theorem (CLT) (Lamperti, 1996, p. 95). Let $x_1, x_2, \ldots$ be i.i.d. random variables having probability distribution $F(x_i)$. Assume that their expected value (μ) and variance (σ^2) are both finite. Define $\bar{x}_n \equiv (x_1 + \ldots + x_n)/n$. Then

[6] See Capiński and Kopp (2004, p. 46) on probability spaces. The technical details here are not relevant to our analysis, but are needed in order to be precise and avoid confusion when invoking widely cited large number results.

(A3) $$P\left(\frac{(\bar{x}_n - \mu)n^{0.5}}{\sigma} \le x\right) \xrightarrow{dist} \mathrm{Nor}_{(0,1)}(x),$$

where $\mathrm{Nor}_{(\mu,\sigma^2)}(x)$ is the cumulative normal distribution with mean μ and variance σ^2.

Comparing the law and the theorem, they may appear to be inconsistent. The SLLN asserts that the distribution of the mean converges to a distribution with discontinuity at μ, or

(A4) $$P\left(\lim_{n\to\infty} \bar{x}_n \le x\right) = \begin{cases} 0, & x < \mu, \\ 1, & x \ge \mu. \end{cases}$$

However, divide through by $n^{0.5}$ and use scaling properties of the normal distribution to consider an alternative version of (A3) as

(A5) $$P\left(\frac{(\bar{x}_n - \mu)}{\sigma} \le x\right) \xrightarrow{dist} \mathrm{Nor}_{(0,n^{-1})}(x).$$

In this case, as $n \to \infty$, then the variance of the mean contracts toward 0, or $\mathrm{Lim}_{n\to\infty}\sigma^2/n = 0$, and the distribution does indeed become degenerate with a discontinuity at $x = \mu$. Factor $n^{0.5}$ in (A3) is required to ensure a nondegenerate limiting distribution. It is often called a "stabilizing" or "blow-up" transformation (Greene, 2003, p. 908).[7]

■ *Proof of Proposition 1.* From the SLLN, we know that the weather-conditioned mean yield over the region converges to $\mu(w)$ with probability 1, or

(A6) $$P\left(\lim_{n\to\infty} \hat{y}_n(w) = \mu(w)\right) = 1, \quad \hat{y}_n(w) = n^{-1}\sum_{i=1}^{n} y_i(w).$$

We will proceed in two steps. The first uses the SLLN to get a weather-conditioned limiting distribution. The second uses the bounded convergence theorem to extend an implication for the unconditional distribution.

Step 1. (Using SLLN). Consider function sequence

(A7) $$u_n(w,y) = \int_Y \int_Y \cdots \int_Y I\left[\hat{y}_n(w) \le y\right] \prod_{i=1}^{n} dF^{|w}(y_i),$$

where $I[C]$ is the indicator function with value 1 if condition C applies, and value 0 otherwise. Here,

$$\int_Y \int_Y \cdots \int_Y$$

represents the n-fold integration over the yields entering the weather-conditioned empirical mean yield in equation (A6).

From almost-sure convergence of $\hat{y}_n(w)$ to $\mu(w)$ under SLLN, it follows that the function limit exists and is given by

(A8) $$u(w,y) \equiv \lim_{n\to\infty} u_n(w,y) = \begin{cases} 1, & \mu(w) \le y, \\ 0, & \mu(w) > y. \end{cases}$$

---

[7] In econometric theory, CLT can be used to transform low variance estimator distributions so that standard significance tests can be employed.

Step 2. (Using the bounded convergence theorem). This theorem, a corollary of the dominated convergence theorem, gives conditions under which the $\text{Lim}_{n\to\infty}$ and $\int$ operations commute.

Theorem (Lewin, 1987). For a bounded interval $A \subset \mathbb{R}$, suppose the sequence of functions $f_n(x)\colon A \to \mathbb{R}$ satisfies

$$\int_A |f_n(x)|\, v(x)\, dx < \infty$$

for $v(x)$ a density function, and for all $n \in \mathbb{N}_0$, $\mathbb{N}_0$ the natural numbers other than 0. Suppose (a) the sequence converges almost everywhere to a limit function $f(x)\colon A \to \mathbb{R}$, and ($b$) there exists a number $\Gamma \in \mathbb{R}$ such that for every $n \in \mathbb{N}_0$, $|f_n(x)|\, v(x) \le \Gamma$ for almost all $x \in A$.[8] Then

$$\int_A |f(x)|\, v(x)\, dx < \infty,$$

and

$$\text{(A9)} \qquad \operatorname*{Lim}_{n\to\infty} \int_A f_n(x) v(x)\, dx = \int_A \operatorname*{Lim}_{n\to\infty} f_n(x) v(x)\, dx = \int_A f(x) v(x)\, dx.$$

In particular, our use of the theorem involves the limiting behavior of $\int u_n(w,y) g(w)\, dw$ since that provides a description of the limiting yield distribution. From (A7), note that

$$\text{(A10)} \qquad \int_W u_n(w,y) g(w)\, dw = \int_W \int_Y \int_Y \cdots \int_Y I\left[\hat{y}_n(w) \le y\right] \prod_{i=1}^{n} dF^{|w}(y_i) g(w)\, dw = P(\bar{y}_n \le y),$$

or the unconditional probability that the mean upon drawing from n plots is no more than yield value y. Taking the limit, noting that $|u_n|\, g(w) \le \Gamma$ to satisfy the theorem's condition (b), and using step 1 to satisfy the theorem's condition (a), the bounded convergence theorem implies

$$\text{(A11)} \qquad \operatorname*{Lim}_{n\to\infty} \int_W u_n(w,y) g(w)\, dw = \int_W u(w,y) g(w)\, dw = \int_{\mu(w)\le y} g(w)\, dw = G\left[\mu^{-1}(y)\right].$$

But $T(y) = G[\mu^{-1}(y)]$. □

■ *Details for Example 1.* When w is normally distributed with mean θ and variance σ^2, then

$$\text{(A12)} \qquad \mathrm{E}\left[e^{tw}\right] = e^{t\theta + 0.5t^2\sigma^2}.$$

Therefore,

$$\text{(A13)} \qquad \begin{aligned} \mathrm{E}[(\mu(w) - \mathrm{E}[\mu(w)])^3] &= \mathrm{E}\left[\left(\lambda_0 - \lambda_1 e^{-\lambda_2 w} - \mathrm{E}\left[\lambda_0 - \lambda_1 e^{-\lambda_2 w}\right]\right)^3\right] \\ &= -\lambda_1^3 \mathrm{E}\left[\left(e^{-\lambda_2 w} - \mathrm{E}\left[e^{-\lambda_2 w}\right]\right)^3\right] \\ &= 3\lambda_1^3 \mathrm{E}\left[e^{-2\lambda_2 w}\right]\mathrm{E}\left[e^{-\lambda_2 w}\right] - 2\lambda_1^3\left(\mathrm{E}\left[e^{-\lambda_2 w}\right]\right)^3 - \lambda_1^3 \mathrm{E}\left[e^{-3\lambda_2 w}\right] \\ &= 3\lambda_1^3 e^{-2\lambda_2\theta + 2\lambda_2^2\sigma^2} e^{-\lambda_2\theta + 0.5\lambda_2^2\sigma^2} - 2\lambda_1^3 e^{-3\lambda_2\theta + 1.5\lambda_2^2\sigma^2} - \lambda_1^3 e^{-3\lambda_2\theta + 4.5\lambda_2^2\sigma^2} \\ &= \lambda_1^3 e^{-3\lambda_2\theta + 1.5\lambda_2^2\sigma^2}\left(3e^{\lambda_2^2\sigma^2} - 2 - e^{3\lambda_2^2\sigma^2}\right). \end{aligned}$$

Also,

[8] Here, "almost all" means almost everywhere or with probability 1. The technical assumption is innocuous in our context.

(A14)
$$\begin{aligned}\mathrm{E}\big[(\mu(w)-\mathrm{E}[\mu(w)])^2\big] &= \mathrm{E}\Big[\big(\lambda_0-\lambda_1 e^{-\lambda_2 w}-\mathrm{E}\big[\lambda_0-\lambda_1 e^{-\lambda_2 w}\big]\big)^2\Big]\\ &= \lambda_1^2\mathrm{E}\Big[\big(e^{-\lambda_2 w}-\mathrm{E}\big[e^{-\lambda_2 w}\big]\big)^2\Big] = \lambda_1^2\mathrm{E}\big[e^{-2\lambda_2 w}\big]-\lambda_1^2\big(\mathrm{E}\big[e^{-\lambda_2 w}\big]\big)^2\\ &= \lambda_1^2 e^{-2\lambda_2\theta+2\lambda_2^2\sigma^2}-\lambda_1^2 e^{-2\lambda_2\theta+\lambda_2^2\sigma^2} = \lambda_1^2 e^{-2\lambda_2\theta+\lambda_2^2\sigma^2}\left(e^{\lambda_2^2\sigma^2}-1\right).\end{aligned}$$

So

(A15)
$$\Upsilon[\mu(w)] = \frac{\lambda_1^3 e^{-3\lambda_2\theta+1.5\lambda_2^2\sigma^2}\left(3e^{\lambda_2^2\sigma^2}-2-e^{3\lambda_2^2\sigma^2}\right)}{\left\{\lambda_1^2 e^{-2\lambda_2\theta+\lambda_2^2\sigma^2}\left(e^{\lambda_2^2\sigma^2}-1\right)\right\}^{3/2}} = \frac{3e^{\lambda_2^2\sigma^2}-2-e^{3\lambda_2^2\sigma^2}}{\left(e^{\lambda_2^2\sigma^2}-1\right)^{3/2}}.$$

Set $p=e^{\lambda_2^2\sigma^2}$ and write this as

(A16)
$$\begin{aligned}\Upsilon[\mu(w)] &= -\frac{(p^3-3p+2)}{(p-1)^{3/2}} = -\frac{(p-1)^2(p+2)}{(p-1)^{3/2}} = -(p-1)^{0.5}(p+2)\\ &= -\left(e^{\lambda_2^2\sigma^2}-1\right)^{0.5}\left(e^{\lambda_2^2\sigma^2}+2\right) < 0.\end{aligned}$$

■ *Proof of Proposition 3.* For parts (i) and (ii), simply extend the proof of proposition 1 to allow for conditioning on z also. For part (iii), note that

(A17)
$$L^{|w}(y) = M\big[r(y;w)\big] = \text{Prob}(\text{yield} \le y \,|\, w),$$

where w has distribution $G(w)$. Then integrate through to obtain the unconditional yield distribution:

(A18)
$$\int_W M\big[r(y;w)\big]\,dG(w) = \int_W \text{Prob}(\text{yield} \le y \,|\, w)\,dG(w) = \text{Prob}(\text{yield} \le y).$$

Similarly, $T^{|z}(y) = G[h(y;z)] = \text{Prob}(\text{yield} \le y \,|\, z)$ and

(A19)
$$\int_Z G\big[h(y;z)\big]\,dF(z) = \int_Z \text{Prob}(\text{yield} \le y \,|\, z)\,dF(z) = \text{Prob}(\text{yield} \le y). \qquad \square$$