On the Distribution of Crop Yields: Does the Central Limit Theorem Apply?

Abstract

In this paper we take issue with the applicability of the central limit theorem (CLT) on aggregate crop yields. We argue that even after correcting for the effects of spatial dependence, systemic heterogeneities and risk factors, aggregation does not necessarily lead to normality. We show that aggregation is also likely to lead to nonnormal distributions, which exhibit both skewness and excess kurtosis. In particular, we consider the case in which the number of summands is not constant but varies with time, which corresponds to the empirically relevant situation where the number of acres used for cultivation of a particular crop exhibits substantial variation over time. In this case, the CLT is not applicable while the limit theorems for random sums of random variables, which apply, predict that the limiting distribution of the sum is not normal and depends on the postulated distribution of the number of summands. Using data from aggregate US states crop yields, we provide empirical support regarding the deviation of aggregate crops yields from normality.

JEL Classification: C16, C51, Q14

Keywords: Aggregate Crop Yield; Central Limit Theorem; Limit Theorem for Random Sums of Random Variables; Normality.

The probability distribution of crop yields has been extensively investigated over the last twenty years or so; however its characterization still remains an open issue. Several authors, such as Just and Weninger (1999), Ker and Goodwin (2000), Atwood, Shaik and Whatts (2002, 2003), Sherrick et al. (2004), Hennessy (2009), to name a few of the recent contributors to this literature, focus on the question of whether crop yields deviate from normality.

Just and Weninger (1999) identify the following reasons, which are likely to prevent the emergence of a general consensus on the shape of crop yield distribution: (i) The lack of a clear pattern against normality. In spite of the presence of ample empirical evidence against normality, the origins of the latter are not at all clear. For example, in an early study, Day (1965) reports weak evidence for positive skeweness and nonnormal kurtosis (both leptokurtosis and platykurtosis) for Mississipi cotton, corn and oats yields. Using aggregate time series data, Gallagher (1987) finds negative skewness in US soybean yields, a result consistent with Taylor (1990). The latter study, however, reports evidence on positive skewness for the wheat yields together with leptokurtosis for all crops (corn soybean and wheat) under consideration (see also Buccola 1986, Moss and Shonkwiler 1993). (ii) The uncertainty surrounding the specification of the conditional mean and conditional variance of yields. Misspecification of the systematic components of crop yields are likely to introduce nonstationarity in the random component thus producing erroneous inferences on the distributional properties of the latter. In the same spirit, Hennessy (2009, p. 46) noted that "...when systemic heterogeneities exist in the data under consideration, these will dominate to determine the shape of the yield distribution". He also provided a link between the skewness of aggregate yield and the weather factor skewness. (iii) Misinterpretation of statistical significance. This problem arises in a univariate framework when one fails to combine the various tests for normality (e.g. separate tests for skewness and kurtosis) into a single test to assess significance.

The same problem is also likely to arise in a multivariate framework, if the possible correlations among yields of several crops are ignored. (iv) The use of aggregate time series (ATS) data to represent farm-specific variation. At each point in time, crop yield data are constructed by taking the acreage-weighted average over the sample farms. This averaging operation eliminates the specific probabilistic features of the yields of each individual farm, thus obscuring the production uncertainty characteristics at the farm-level.

All issues raised above are indeed valid. However, Just and Weninger make an additional point concerning the necessity of the normal distribution as the appropriate probabilistic description of crop yields, which arises from the fact that the CLT seems to be at work. Specifically, Just and Weninger correctly point out that "crop yields at all levels are averages". In particular, they state: "At the aggregate level, ATS data are averages of yields over many farms. At the farm-level, yields are averages of production over many acres" (pp. 301). As a result, the above mentioned authors conclude that "under broad conditions" the probability distribution of these averages has to be the normal because of the CLT. On the other hand, Goodwin and Ker (1998) and Goodwin and Mahul (2004), among others, state that the existence of spatial dependence and systemic risk factors indicate that a straightforward application of the CLT is not appropriate.

In this paper we take issue with the applicability of the CLT on aggregate crop yields, arguing that even after correcting for the effects of spatial dependence, systemic heterogeneities and (systemic) risk factors, aggregation does not necessarily lead to normality but instead it is also likely to lead to nonnormal distributions, which exhibit both skewness and excess kurtosis. Put it differently, although we do accept the fact that crop yields are indeed averages and also that "convergence in distribution" seems to be in place (in the sense that no distributional explosion is observed) we do not accept that the only possible limiting distribution is the normal. More specifically, we consider the case in which the number of summands is not constant but varies with time, being a random variable itself. This corresponds to the case in which the number of acres used for the cultivation of a particular crop exhibits substantial variation over time. Indeed, one of the most critical decisions that a farmer makes is what crops to grow on the land she has available. For some farmers, the decision of which crops to cultivate is straight forward because the land, climate, tradition, infrastructure and economic conditions all support one dominant crop[1]. However, these farmers still need to decide each year how many acres of this crop they will cultivate, which may be a function of unpredictable economic conditions. For other farmers, there may be a variety of crops adapted to their local ecology, and they may wrestle each year with the decision of what crops to plant on what pieces of land. Factors that can affect cropping decisions in a random way are predictions about the weather and predictions on what crops may be planted in other parts of the country or the world which will influence expectations about prices for different crops at the end of the growing season.

Under this set of assumptions, the central limit theorem is not applicable; instead we must appeal to limit theorems for random sums of random variables (see Gnedenko and Korolev 1996). These theorems predict that the limiting distribution of the sum is not normal and depends on the postulated distribution for the number of summands (see Clark 1973 and Blattberg and Gonedes 1974, among others, for an application of these ideas to stock returns). Using data from US aggregate state crop yields, we provide empirical support for the predictions of these theorems. In particular, we find positive correlation between crop-specific acreage and a set of statistics that measure the deviation of aggregate crops yield from normality.

This paper is organized as follows: first, the case against standard convergence to normality mentioned above is analyzed in detail, then the relevant empirical support is provided using US data. Last section concludes the paper.

On the Applicability of the Central Limit Theorem: The case of Random Sums

As mentioned in Introduction, Just and Weninger (1999) make a case for the normality of the distribution of crop yields by appealing to the CLT. More specifically, they claim that "At the aggregate level ATS data are averages of yields over many farms. At the farm-level, yields are averages of production over many acres. The CLT implies that averages have asymptotically normal distributions under broad conditions" (Just and Weninger 1999, pp 301). Let us formalize this statement. First, we may assume that the random variable of interest is the production of a specific acre, with corresponding index j, at time t, denoted by ξ_{jt}. Obviously, the values of the random variable ξ_{jt} (that is the production of each specific acre j at time t) are not observable. Nevertheless, what is observable is the production at the State, county, or even farm level as well as the total acreage of each State (or county or farm) devoted to the production of a specific crop. So, let X_{it} and n_{it} be the production and the acreage of farm i, respectively, k_t be the number of farms and $N_t = \sum_{i=1}^{k_t} n_{it}$ be the total acreage of all farms at time t. Then, the average yield per acre is given by:

$$\overline{y}_t = \frac{\sum_{i=1}^{k_t} X_{it}}{\sum_{i=1}^{k_t} n_{it}} = \frac{\sum_{j=1}^{N_t} \xi_{jt}}{N_t} \tag{1}$$

More specifically, we must distinguish between the probabilistic properties of the ξ_{jt}'s within the same time period t (cross-sectional properties) and those across time (temporal properties). To formalize this, we may arrange the random variables ξ_{jt} into the following array:

$$
\begin{array}{lccc}
 & t=1 & t=2 & ... \\
j=1 & \xi_{11} & \xi_{12} & \cdots \\
j=2 & \xi_{21} & \xi_{22} & \cdots \\
. & & & \\
. & & & \\
. & & & \\
j=N_t & \xi_{N_1 1} & \xi_{N_2 2} & \cdots
\end{array}
\tag{2}
$$

where the last line in (2) does not correspond to a specific row but it describes the last element of each column.

In this general setting, the random variables ξ_{jt} may be characterised by two types of dependence. The first one is cross-sectional dependence, that is dependence among the elements of the columns of (2). The second is temporal dependence among the elements of the rows of (2). Put it differently, cross sectional dependence refers to dependence among the crop yields of various acres within the same time period, whereas temporal dependence concerns the dependence among ξ_{jt}s across different time periods. Similar distinctions can be made about time and spatial heterogeneity. More specifically, cross-sectional heterogeneity concerns the extent to which the distributions, $D(\cdot)$, of the yields of various acres within the same period are different, whereas time heterogeneity refers to the distributions of the total yield across different time periods. For example, if $D(\xi_{11}) = D(\xi_{21}) = ... = D(\xi_{N_1 1})$ then we have cross-sectional homogeneity for $t = 1$. On the other hand, if $D(\overline{y}_1) = D(\overline{y}_2) = ...$ then the distributions of the average yield over time are identical.

Let μ_t and σ_t denote the mean and standard deviation of ξ_{jt} at time t. In other words, we assume that the first two moments of ξ_{jt} are equal across $j = 1, 2, ..., N_t$ (cross-sectional homogeneity of the second-order). The CLT states that, under some additional conditions on the probabilistic properties of the individual random variables ξ_{jt},

$$
\frac{\sqrt{N_t}\,(\overline{y}_t - \mu_t)}{\sigma_t} \xrightarrow{L} N(0,1). \tag{3}
$$

As an implication of (3), we have that for large N_t,

$$\overline{y}_t \simeq N(\mu_t, \sigma_t^2/N_t) \tag{4}$$

Remarks

(i) Just and Weninger claim that the (cross-sectional) conditions under which (4) is true, are "broad". Indeed, recent results in Probability Theory on the conditions under which CLT applies seem to make a very strong case in favor of the approximate normality of $\overline{y}_t$. More specifically, it has been proved that CLT holds under quite general properties for the initial sequence $\{\xi_{jt}\}_{j\geq 1}$ (with fixed t). For example, Ibragimov (1962) proves that $\{\xi_{jt}\}_{j\geq 1}$ obeys CLT if it is strictly stationary, $\alpha-$mixing sequence with $E\left|\xi_{1t}\right|^{2+\delta_t} < \infty$, for some $\delta_t > 0$. Herrndorf (1984) relaxes the assumption of stationarity and derives a CLT for $\alpha-$mixing sequences of random variables satisfying the condition $\sup_{i\in\mathbb{N}} E\left|\xi_{it}\right|^{b_t} < \infty$, for some $b_t > 2$ (see Kourogenis and Pittis 2009 for a survey of CLT's). On the contrary, when the area under consideration is restricted, then spatial dependence may not dissipate fast enough for CLT to hold (see Goodwin and Ker (1998), Goodwin and Mahul (2004) among others).

(ii) It is obvious from (4) that a sufficient condition for achieving time homogeneity amounts to $\mu_t = \mu$ and $\sigma_t = \sigma^2 > 0$ for every t.

One of the assumptions implicit in (4) is that the number, N_t, of summands is large and "certain". In many interesting cases, however, the number of the summands is not constant but is itself a random variable. In such cases, it is interesting to investigate the limiting behaviour of the so-called "random sums" of random variables. More specifically, we are interested in finding the conditions (if any) under which $N(0,1)$ is still a good approximation of the distribution of the aggregate crop yields, if the number, N_t (the total acreage) of the ξ_{jt}'s (production per acre) is large but random. Put it differently, we are interested in examining whether there are any conditions under which, for each t, the ξ_{jt}'s may still belong to the domain of attraction of the normal law, even in the presence of randomness in the number of summands. We are also interested in identifying the cases for which a distribution, D, different than $N(0,1)$ is the appropriate limiting distribution of the random sum and studying its properties.

To define the problem, for each t, let $\{\xi_j\}_{i\geq 1}$ be an iid sequence of random variables with finite $E(\xi_j) = \mu_\xi$ and $Var(\xi_j) = \sigma_\xi^2 > 0$. Obviously, the moments μ_ξ and σ_ξ^2 may in general vary across time, but since the analysis that follows refers to a specific time period, we choose to drop the second subscript t from the more appropriate notation μ_{ξ_t} and $\sigma_{\xi_t}^2$ for simplicity. Moreover, (and following the same notational convention of dropping the time subscript) let $\{N_n\}_{n\geq 1}$ be a sequence of nonnegative integer-valued random variables. Define the random sum process

$$S_{N_n} = \sum_{i=1}^{N_n} \xi_j.$$

The question to be answered is under what conditions the random sum, S_{N_n}, properly normed and centered, converges in law to some random variable, Z, and further, under what additional conditions Z is distributed as $N(0,1)$. Robbins (1948) obtains sufficient conditions for the convergence in law of the properly centered and normed sequence S_{N_n}, to normal, under the assumption that N_n is independent of the summands, ξ_1, ξ_2,.... Renyi (1960) and Blum, Hanson and Rosenblatt (1963) derive sufficient conditions similar to those of Robbins (1948) without the assumption of independence between N_n and the summands.

To state the problem formally, (for each t) define the centered and normed random sum process, Z_n, as follows:

$$Z_n = \frac{S_{N_n} - n\mu_\xi}{\sqrt{n\sigma_\xi^2}}$$

We are interested in finding the general conditions under which the sequence of the Z_n's converges in law to a random variable Z, as well as the specific conditions guaranteeing that $Z \sim N(0,1)$. Finkelstein and Tucker (1989) show that under the assumption that N_n is independent of the summands, the necessary and sufficient condition for

$$Z_n \xrightarrow{L} \text{(some) } Z \tag{5}$$

is given by

$$U_n \equiv \frac{N_n - n}{\sqrt{n}} \xrightarrow{L} \text{(some) } U. \tag{6}$$

In such a case, the distribution of Z is that of the sum of two independent random variables, Z_1 and Z_2, where Z_1 is $N(0,1)$ and the distribution of Z_2 is the same to that of $\frac{\mu_\xi U}{\sigma_\xi}$. This result may be stated in an equivalent way by saying that the distribution function F_Z of Z is equal to $\Phi * F_U$ where Φ and F_U are the distribution functions of the standard normal and U, respectively. In other words, Z is a mixture of normals, with the mean being mixed by U (see Finkelstein, Kruglov and Tucker 1994).

Finkelstein and Tucker (1989) also derive the conditions for the convergence of Z_n to $N(0,1)$. Specifically, they show that if (and only if) condition (6) is replaced by

$$U_n \equiv \frac{N_n - n}{\sqrt{n}} \xrightarrow{p} 0, \tag{7}$$

then

$$Z_n \xrightarrow{L} N(0,1). \tag{8}$$

Remarks:

(i) Condition (7) is stronger than (6). This in turn implies that the assumptions that must be made on the behavior of $\{N_n\}_{n\geq 1}$ in order to obtain asymptotic normality are stronger than the ones that ensure simply convergence to some distribution. In the case that (6) holds but (7) fails, the sequence Z_n of random sums converges in law to a non-normal random variable, which is likely to exhibit both skewness and excess kurtosis.

(ii) The case analyzed above is usually referred to as "convergence of random sums under nonrandom centering". This is due to the fact that the random sum process S_{N_n} is centered by the sequence of constants, $n\mu_\xi$. A somewhat different problem arises in the case that the sequence S_{N_n} is centered by $N_n\mu_\xi$ instead of $n\mu_\xi$. In such a case, the random sum process is centered by a sequence of random variables rather than by a sequence of constants. This asymptotic problem, referred to as "convergence of random sums under random centering" was first analyzed by Renyi (1960) who showed that the centered and normed random sum process,

$$Z_n^* \equiv \frac{S_{N_n} - N_n\mu_\xi}{\sigma_x\sqrt{n}} \tag{9}$$

converges to $N(0,1)$ if

$$\frac{N_n}{n} \xrightarrow{p} 1. \tag{10}$$

Condition (10) which ensures convergence to $N(0,1)$ in the case of random centering is weaker and hence easier to be satisfied in practice than (7) which corresponds to the case of nonrandom centering. However, nonrandom centering is more natural for constructing approximate distributions. Moreover, Korolev (1995) argues that random centering "significantly restricts the class of possible limit laws compared to the general situation, where random sums are centered by constants" (1995, pp. 2153).

(ii) The results analyzed above were obtained under the restrictive assumption that the initial random variables, ξ_j, are iid with finite mean, μ_ξ, and variance, σ_ξ^2. Billingsley (1962) extends this result by proving the asymptotic normality of Z_n^* for the cases in which (10) holds and the initial sequence $\{\xi_j\}_{i\geq 1}$ obeys the "invariance principle" (IP). The latter is a stronger version of the classical central limit theorem for nonrandom partial sums. It has been shown that IP is satisfied by a wide spectrum of non iid sequences $\{\xi_j\}_{i\geq 1}$, such as strong or uniform mixing ones (see Kourogenis and Pittis 2009 for a recent survey on this topic). This result implies that $Z_n^* \xrightarrow{L} N(0,1)$ holds under (10) even for cases in which $\{\xi_j\}_{i\geq 1}$ is an asymptotically independent, nonstationary sequence.

The practical implications of the preceding discussion may be summarized as follows:

(i) When a random variable, Z, is the sum of elementary random variables, then its distribution may be approximated by the normal one, even if the number of summands, N_n, is random. This is valid when N_n behaves in a way prescribed by conditions (7) or (10). Specifically, N_n must exhibit small variation around n for large n. If N_n displays considerable variability around n even for large n, then the asymptotic distribution of Z is not normal but rather a mixture of normals. In such a case, the empirical distribution of Z is likely to exhibit both skewness and excess kurtosis.

(ii) In assessing the distribution of crop yields using aggregate time series data, we face the following problems: First we must account for possible trends in the aggregate series arising from time heterogeneity in the moments of the ξ_j's. More specifically, if we assume that $E(\xi_{jt}) = \mu_\xi t$ then the aggregate crop yield series will exhibit a trending behavior, which has to be accounted for before any tests for normality are carried out. This issue is analyzed in the third and fourth sections of Just and Weninger (1999) and is also considered in the empirical section of this paper. However, even if we succeed in correctly detrending the aggregate series, we still face the problem of the possible variation (randomness) of the number of acres that enter the calculation of the aggregate yield over time. If this variation is substantial (in the sense that it violates conditions (7) or (10)), then non-normality of the aggregate data is likely to arise.

Verbally, we consider the case in which the number of summands is not constant but varies with time, being a random variable itself. This corresponds to the case in which the number of acres used for the cultivation of a particular crop exhibits substantial variation over time and this variation is more or less random. As discussed in Introduction, factors that can affect cropping decisions in a random way are predictions about the weather and predictions on what crops may be planted in other parts of the country or the world which will influence expectations about prices for different crops at the end of the growing season.

Empirical Results

The analysis of the previous section suggests that the presence of non-normality in the crop yield distributions is likely to derive from the random nature of the number of acres employed in the production of various crops over time. This assumption implies that we should observe some significant correlation between the sample standard deviation, $s(\Delta N_t)$, of the percent annual changes, ΔN_t, of the total number of acres employed in the production of a specific crop and any measure of non-normality (such as skeweness and excess kurtosis coefficients) of the distribution of the aggregate (State level) yield of this crop. To examine this empirical implication, we first estimate the skeweness, α_3, and kurtosis, α_4, coefficients of the distribution of percent annual changes, Δy_t, of the yields of five major crops, namely cotton, soybean, corn, barley and wheat together with the value of the Jarque-Berra (JB) test for normality for several US States[2]. We also estimate the same parameters for the residuals, u_t, of an auxiliary autoregression of Δy_t on Δy_{t-1}, Δy_{t-2} and a time trend. The latter case aims at controlling for non-normality effects caused by the presence of temporal dependence and/or time heterogeneity (deterministic or stochastic) in the original crop yield series (see Just and Weninger 1999 for a detailed discussion of these points).

Table 1 around here

Table 1 reports the following correlation coefficients: (i) the correlation between $s(\Delta N_t)$ and the absolute value of α_3, (ii) the correlation between $s(\Delta N_t)$ and the absolute value of $\alpha_4 - 3$, (iii) the correlation between $s(\Delta N_t)$ and JB. Note that the employed distributional characteristics have been calculated for two alternative empirical distributions of crop yields. The first one refers to the raw data of Δy_t whereas the second one corresponds to the residuals u_t. The results may be summarized as follows:

(i) All estimated correlation coefficients have positive sign thus suggesting a positive relationship between the standard deviation of ΔN_t and each of the employed measures of non-normality.

(ii) The magnitude of these correlation coefficients, in general, seems to be higher for the case in which the detrended and demeaned crop yield series are employed. For example, for the case of soybean, the correlation coefficient between $s(\Delta N_t)$ and $|\alpha_3|$ is equal to 0.24 and 0.50 for the cases of raw and filtered crop yield series respectively.

(iii) In some cases, the estimated correlation coefficients exceed the value of 0.5, thus reaching an impressively high value. For example, the correlation coefficient between $s(\Delta N_t)$ and $|\alpha_4 - 3|$ for the case of the filtered cotton yield is equal to 0.54, whereas the same coefficient for the case of filtered barley yield reaches the value of 0.70.

(iv) When the residuals u_t are employed, the smallest correlation coefficient is the one between $s(\Delta N_t)$ and the Jarque-Berra Statistic for the case of wheat and is equal to 0.08. It is interesting to note that this is the crop for which ΔN_t exhibits the smallest average variation across States. More specifically, the mean of the estimated $s(\Delta N_t)$'s across States is equal to 30.71, 36.22, 25.40, 34.35 and 25.1 for cotton, soybean, corn, barley and wheat, respectively. This piece of evidence implies that the minimum "correlation effects" appear in the case of wheat for which the percent annual changes of the total number of acres displays the minimum variation among the five crops under consideration.

Table 2 reports the estimated regression coefficients and the corresponding t-statistics between the standard deviation of ΔN_t (explanatory variable) and the the distributional characteristics of crop yield changes, as measured by the skeweness, the kurtosis and the Jarque-Berra Statistic:

Table 2 around here

In accordance with the results of Table 1, we find that when the statistics of the residuals u_t are used, all crops have at least one regression coefficient with corresponding t-statistic greater than 1.96. More specifically:

(i) The higher values of the t-statistics correspond to barley (all greater than 5) and to corn (all greater than 2.1).

(ii) The absolute values of the t-statistics when u_ts are used are in general higher than the corresponding ones for the case where Δy_ts are employed.

Conclusions

This paper comments on the assertion of Just and Weninger (1999) that the distribution of the aggregate crop yields is expected to be normal due to the applicability of the CLT. We argue that normality is not an inevitable consequence of the operation of aggregation of crop yields. Motivated by the empirical observation that the number of crop-specific acres exhibits substantial variation over time (due to weather predictions or predictions about cultivation decisions elsewhere that will affect expectations on crop prices), we consider limit theorems that are applicable when the number of summands is not constant but varies with time. These theorems predict that the limiting distribution of the sum is not normal and depends on the postulated distribution for the number of summands.

Our empirical analysis investigates the existence of significant correlation between the sample standard deviation of the percent annual changes of the total number of acres employed in the production of a specific crop (ΔN_t) and different measures of non-normality (skeweness and excess kurtosis coefficients) of the distribution of the aggregate yield of this crop. We apply this investigation to five major crops, namely cotton, soybean, corn, barely and wheat. To hedge against the presence of non-normality effects due to temporal dependence and/or time heterogeneity in the original crop yield series, we apply the same investigation to the de-trended and demeaned series of the same crops. Our results provide empirical support for our theoretical predictions. In particular, we find a positive relationship between ΔN_t and different measures of non-normality, the magnitude of which increases, reaching impressively high values for some crops, when we use the de-trended and demeaned crop series.

Our results have implications for the correct specification and estimation of econometric models of crop yields, since we have identified an additional factor, namely the standard deviation of ΔN_t, which can cause nonzero skewness and excess kurtosis in the distribution of aggregate crop yields. This implies that when policy-making is based on these estimated models, one needs to be cautious to take into account the changes in crop-specific acreage, in order to avoid unreliable and misleading results deriving from distributional mispecification.

References

[1]Atwood, J., Shaik, S., Watts, M. (2002). Can Normality of Yields Be Assumed for Crop Insurance? *Canadian Journal of Agricultural Economics* 50, 171-184.

[2]Atwood, J., Shaik, S., Watts, M. (2003). Are Crop Yields Normally Distributed? A Reexamination. *American Journal of Agricultural Economics* 85, 888-901.

[3]Billingsley, P. (1962). Limit Theorems for Randomly Selected Partial Sums. *Annals of Mathematical Statistics* 33, 85-92.

[4]Blattberg, R. C., Gonedes, N. J. (1974). A Comparison of the Stable and Student Distributions as Statistical Models for Stock Prices. *Journal of Business* 47, 244-280.

[5]Blum, J. R., Hanson, D. L., Rosenblatt, J. I. (1963). On the central limit theorem for the sum of a random number of independent random variables. *Probability Theory and Related Fields* 1, 389-393.

[6]Buccola, S. T. (1986). Testing for Nonnormality in Farm Net Returns. *American Journal of Agricultural Economics* 68,. 334-343.

[7]Clark, P. K. (1973). A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices. *Econometrica* 41, 135-155.

[8]Day, R. H. (1965). Probability Distributions of Field Crop Yields. *Journal of Farm Economics* 47, 713-741.

[9]Finkelstein M., Tucker, H. G. (1989). A Necessary and Sufficient Condition for Convergence in Law of Random Sums of Random Variables Under Nonrandom Centering. *Proceedings of the American Mathematical Society* 107, 1061-1070.

[10]Finkelstein, M., Kruglov, V. M., Tucker, H. G. (1994). Convergence in law of random sums with non-random centering. *Journal of Theoretical Probability* 7, 565-598.

[11]Gallagher, P. (1987). U.S. Soybean Yields: Estimation and Forecasting with Nonsymmetric Disturbances. *American Journal of Agricultural Economics* 69, 796-803.

[12]Gnedenko, B. V., Korolev, V. Y. (1996). *Random summation: limit theorems and applications.* CRC Press LCC, Boca Raton, Florida.

[13]Goodwin, B. K., Ker, A. P. (1998). Nonparametric Estimation of Crop Yield Distributions: Implications for Rating Group-Risk Crop Insurance Contracts. *American Journal of Agricultural Economics* 80, 139-153.

[14]Hennessy, D. A. (2009). Crop Yield Skewness and the Normal Distribution. *Journal of Agricultural and Resource Economics* 34, 34-52.

[15]Herrndorf, N. (1984). A functional central limit theorem for weakly dependent sequences of random variables. *Annals of Probability* 12, 141-153.

[16]Ibragimov, I. A. (1962). Some limit theorems for stationary processes. *Theory Prob. Appl.* 7, 349-382.

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

[18]Ker, A. P., Goodwin, B. K. (2000). Nonparametric Estimation of Crop Insurance Rates Revisited. *American Journal of Agricultural Economics* 82, 463-478.

[19]Kourogenis, N., Pittis, N. (forthcoming). "Mixing Conditions, Central Limit Theorems and Invariance Principles: A Survey of the Literature with Some New Results on Heteroscedastic Sequences", *Econometric Reviews.*

[20]Korolev, V. Y. (1995). Limit behavior of centered random sums of independent identically distributed random variables. *Journal of Mathematical Sciences* 76, 2153-2162.

[21]Moss, C. B., Shonkwiler, J. S. (1993). Estimating Yield Distributions with a Stochastic Trend and Nonnormal Errors. *American Journal of Agricultural Economics* 75, 1056-1062.

[22]Renyi, A. (1960). On the central limit theorem for the sum of a random number of independent random variables. *Acta Math. Acad. Sci. Hungar.* 11, 97-102.

[23]Robbins, H. (1948). On the Asymptotic Distribution of the Sum of a Random Number of Random Variables. *Proc. Nat. Acad. Sci.* 34, 162-163.

[24]Sherrick, B. J., Zanini, F. C., Schnitkey, G. D., Irwin, S. H. (2004). Crop Insurance Valuation under Alternative Yield Distributions. *American Journal of Agricultural Economics* 86, 406-419.

[25]Taylor, C. R. (1990). Two Practical Procedures for Estimating Multivariate Nonnormal Probability Density Functions. *American Journal of Agricultural Economics* 72, 210-217.

Footnotes:

1: For example, in the South of the US cotton was king because it grew well in the long, hot summers, the farmers understood how to manage it and the cotton gins, markets and transportation systems were all nearby.

2: Data are collected from National Agricultural Statistics Service (NASS). NASS publishes annual time series on harvested land and yield production for a variety of commodities both in county, state and country level. Selected crops satisfy a minimum requirement of 50 observations (that is collecting data for at least half a century) for harvested land and crop's yields. This condition, depending on the crop examined, resulted in excluding states that did not track down these series for a long period. Therefore, we included in our study 17, 31, 46, 39 and 44 states in the case of cotton, soybean, corn, barley and wheat respectively. The inception year of available data varies from 1866 (143 annual observations) to 1959 (50 annual observations).

Table 1: Correlation Between the Standard Deviation of ΔN_t and the Distributional Characteristics of Crop Yield Changes

		Jarque-Berra Statistic	Skewness	Kurtosis
	Cotton	0.51	0.34	0.53
	Soybean	0.09	0.24	0.14
Δy_t	Corn	0.30	0.25	0.26
	Barley	0.34	0.49	0.43
	Wheat	0.02	0.30	0.11
	Cotton	0.44	0.25	0.54
	Soybean	0.32	0.50	0.31
u_t	Corn	0.36	0.31	0.33
	Barley	0.75	0.66	0.70
	Wheat	0.08	0.32	0.16

Table 2: Estimated Regression Coefficients (t-statistics in parentheses) Between the Standard Deviation of ΔN_t (explanatory variable) and the Distributional Characteristics of Crop Yield Changes.

		Jarque-Berra Statistic	Skewness	Kurtosis
	Cotton	0.10 (2.27)	0.004 (1.40)	0.014 (2.43)
	Soybean	0.11 (0.48)	0.00 (1.35)	0.01 (0.77)
Δy_t	Corn	9.60 (2.11)	0.01 (1.72)	0.05 (1.82)
	Barley	8.09 (1.02)	0.03 (3.12)	0.11 (1.72)
	Wheat	0.48 (0.54)	0.006 (2.13)	0.014 (1, 07)
	Cotton	0.04 (1.90)	0.002 (0.98)	0.013 (2.47)
	Soybean	0.30 (1.81)	0.01 (3.09)	0.02 (1.76)
u_t	Corn	9.90 (2.58)	0.01 (2.19)	0.06 (2.30)
	Barley	8.81 (6.94)	0.02 (5.34)	0.09 (6.01)
	Wheat	0.19 (0.14)	0.006 (1.98)	0.012 (0.71)