Measurement of Enterprise Variability by the Variate Difference Method

By Gerald W. Dean and Harold O. Carter

Much has been written concerning the importance of risk and uncertainty on decision making. However, research results employing static theory are seldom modified by risk and uncertainty considerations to provide more realistic recommendations to farmers and others making decisions under imperfect knowledge. Too often, for example, farm plans derived by budgeting or linear programming are unqualified recommended as "optimum" because they provide maximum profits under average or "normal" prices and yields. To make such results more meaningful, the farmer also needs some estimate of the risk or uncertainty associated with the plans. Ordinarily, the farmer's view of this uncertainty is highly subjective since his past experiences are often limited (that is, in the case of new farmers) or based on a "biased" sample of years. Thus, farmers need a more objective measurement of the uncertainty or variability associated with various enterprises and combinations of enterprises. Our contribution concerning this problem is published in the Giannini Foundation Paper series. The authors offer acknowledgment to C. O. McCorkle, Jr., G. M. Kuznets, and G. Tintner for helpful suggestions in various phases of the study and preparation of the manuscript.

THE purpose of this paper is to indicate the possibilities, advantages, and limitations of the variate difference method for estimating variability measures for individual crops and cropping combinations. Variability measures of this type can be used effectively by teaching and extension personnel. In addition, such information may be incorporated by researchers into certain linear programming problems.[1] We shall argue that the variate difference method [2] may more nearly isolate the truly "random" or "unpredictable" component of total variability than alternative methods and hence provide a more relevant measure of risk or uncertainty.[3] Finally, empirical applications of the variate difference method for estimating crop production variability in California are briefly considered.

Measurement of Variability

In the preceding paragraph, the terms "risk" and "uncertainty" are used loosely to characterize the general framework of imperfect knowledge within which decision makers operate. More precisely, following Knight,[4] risk situations are

---

[1] The variance and covariance measures derived by the method investigated in this paper appear to be applicable in risk or stochastic linear programming studies. The authors are currently investigating this possibility in more detail. For previous work in this area, see: Babbar, M. M., "Distribution of Solutions of a Set of Linear Equations," Jour. Amer. Statis. Assoc., Vol. 50, 1955, p. 854; Tintner, G., "Stochastic Linear Programming," Second Symposium on Linear Programming, U.S. Bureau of Standards, Washington, D.C., 1955, p. 197; Freund, R. J., "The Introduction of Risk into a Programming Model," Econometrica, Vol. 24, 1953, p. 253; Heady, E. O. and Candler, Wilfred, *Linear Programming Methods*, The Iowa State College Press, Ames, Iowa, 1958, p. 554.

[2] Tintner, Gerhard, *The Variate Difference Method*, Bloomington, Ind.: Principia Press, Inc., 1940 (Cowles Commission for Research in Economics Monograph No. 5), 175 pp.

[3] Heady, Brown, Botts, and Kling have derived similar measures of variability with different methods for both crops and livestock production based on time series data. See for example: Heady, E. O., Kehrberg, E. W., and Jebe, E. H., *Economic Instability and Choices Involving Income and Risk in Primary or Crop Production,* Iowa Agr. Expt. Sta. Res. Bul. 404, 1954; Brown, W. G. and Heady, E. O., *Economic Instability and Choices Involving Income and Risk in Livestock and Poultry Production,* Iowa Agr. Expt. Sta. Res. Bul. 431, 1955; Botts, R. R., *Variability of Cotton Yields,* U.S. Bur. Agr. Econ. August, 1952 (Mimeo); Botts, R. R. and Barber, E. L., *Variability of Corn Yields,* U.S. Bur. Agr. Econ., July 1952, (Mimeo); and Kling, William, "Determination of Relative Risks Involved in Growing Truck Crops," Jour. Farm Econ. Vol. 24, August, 1942, No. 3.

[4] Knight, Frank H., *Risk, Uncertainty and Profit,* Boston: Houghton Mifflin Co., 1921.

those in which parameters (such as the mean and variance) of the probability distribution of outcomes can be established empirically; uncertainty situations are those in which such parameters cannot be objectively established. Thus, any study that attempts to estimate empirically "measures of variability" falls more nearly in the classical *risk* setting. In fact, by providing objective measures of the variability of outcomes, the researcher is attempting to transfer decision makers from an *uncertainty* to a *risk* setting. Thus, rather than relying on a "subjective guess," the decision-maker receives quantitative estimates of variability to guide his actions.

Variability in agriculture stems from the fact that crop yields, livestock gains per feed unit, prices, costs, and incomes are influenced by many variables—some in a systematic or rather "predictable" fashion and others in an unpredictable or "random" manner, at least when viewed *ex ante*. Imperfect knowledge of the future stems primarily from the random or unpredictable component. While, by definition, the value of a random component in any one year cannot be predicted, parameters (such as the variance) of the distribution of the random component might be estimated as guides in decision making. However, a difficult question arises: From the standpoint of the individual farmer, what portion of total variability is really unpredictable or random and what portion is predictable?[5]

The most naive assumption is that any deviation from the long-run mean is a random or unpredictable event in the eyes of the farmer. Such a procedure essentially represents a "no knowledge" situation. More realistically, farmers probably recognize certain long-run physical and economic trends over time, such as the advancing level of technology, inflation, and price cycles. For example, farmers planning crop production for the year ahead are more likely to view the random element of yields or prices as a deviation from the "current level" rather than as a deviation from the long-run mean.[6]

Several different empirical procedures are available for determining the exact current level of the time series (and hence for determining the deviations from this current level). One familiar technique is to approximate the current level of the time series by a fitted trend line, then to assume that deviations from trend represent the random component.[7] A second method is to assume that the current level is identical with the observation in the previous year. In this case, the random element is identical with first differences of the data.[8] A third procedure might be to approximate the current level by a moving average, then to assume that deviations from this moving average constitute the random element. A price series might be deflated by some general price index to arrive at "real" values of the series, then the deviations from the long-run mean of the deflated series would be assumed to represent the random element.

Arguments for and against each of these procedures might be advanced. But the one that seems to the authors to be most reasonable is the first method of trend removal. Even this procedure is based on the limiting assumption that the systematic component of the time series (that is, the general price level, technological trend, and so on) can be characterized by linear, polynomial, or other types of mathematical functions. The authors prefer a statistical method that does not depend on rigid functions which may be difficult to defend on economic grounds. The variate difference method seems to meet this objection.

The fundamental assumption of the variate difference method is that every economic time series consists essentially of two additive parts. The first part is the mathematical expectation or systematic component of the time series in which consecutive observations are positively correlated with each other. This does not imply that the procedure is restricted to series showing a positive trend, that is, a negatively sloped line also produces positively correlated consecutive observations. However, the method is inappropriate for

---

[5] It is recognized that certain fluctuations which might be classed as unpredictable or "random" to the individual farmer could in fact be "explained" by appropriate aggregate supply and price analysis.

[6] By "current" level is meant the general level at the time the decision is made, rather than the level in 1959.

Thus, with respect to any past year, the "current" level refers to the general level prevailing in that particular year.

[7] Crop yields have been handled in this way by Heady, et al., *op. cit.*, p. 627.

[8] See: Kling, *op. cit.*, p. 695.

excessively "zigzag" series. The second part is the random or unpredictable component in which consecutive items are assumed not to be auto-correlated.[9] The variate difference method is appropriate for separating out the random portion of time series because it avoids unnecessary assumptions about the functional character of the systematic component. It is assumed that the smooth part of the time series (the systematic component) can be approximated by polynomials of the variable time which otherwise need not be specified. A well-known theorem regarding a polynomial of degree m is that its m-th finite difference is constant and its $m+1$, $m+2$, . . . finite differences vanish. However, the random component cannot be reduced by finite differencing since it is not ordered in time (that is, it cannot be approximated by a smooth function). Thus, the method is designed to eliminate the systematic component by successive finite differencing, leaving an estimate of the random element.

Estimating Individual Enterprise Variability by Variate Difference Method

Using the variate difference method, the variance of a time series can be "split up" into two parts one of which comes from the mathematical expectation and the second of which is the variance of the random component. As indicated above, interest is in this latter quantity as an estimate of enterprise variability. The method consists of calculating variances of the original series and of the series of successive finite differenes. If a finite difference of the order k_0 can be found such that the variance of the k_0^{th} difference is equal to the variance of the $(k_0+1)^{th}$ difference and equal to that of the $(k_0+2)^{th}$ difference, and so on, it is reasonable to assume that the mathematical expectation has been eliminated to a reasonable degree by taking k_0 differences.[10] The difference between the variances of two successive series of finite differences is compared with

its standard error in order to decide when the variances are approximately equal. If the difference is smaller than about three times its standard error, it is reasonably certain that, from a probability viewpoint the finite differencing has been carried sufficiently far to have eliminated the nonrandom element; remaining is an estimate of the variance of the random component of the time series.

As an example of the method, table 1 summarizes the computations used in estimating the variance of the random component of early fall lettuce yields in California for the years 1918–57. The question is: Beginning with which difference is it reasonably certain that the nonrandom element has been eliminated, leaving an estimate of the random variance? Table 1 indicates that the variance does not stabilize until the second difference, as shown by the standard error ratios of 5.44, 4.23, and 1.09. Therefore, 224.72 is taken as the estimate of the variance of the random component.

Empirical Results

Table 2 indicates the relative variability in yield, price, and gross income of selected California crops using the variate difference method.[11] The variability coefficient (equation 1) expresses the square root of the random variance (standard deviation of the random component) as a percentage of the 1953–57 mean of the series;[12]

(1) Variability coefficient $= \dfrac{\sqrt{\text{random variance}}}{1953\text{–}57 \text{ mean}} \times 100$, variability relative to recent levels seems

---

[9] It should be emphasized that the variate difference method is only applicable if there is no auto-correlation in the random element. See Tintner, Gerhard, *Econometrics.* John Wiley and Sons, Inc., New York, 1952, pp. 312–314.

[10] Tintner, *op. cit.*, p. 33.

[11] While net income variability is of ultimate interest to farmers, lack of adequate cost data prevented its derivation for all crops. Net income variability is used, however, where crop combinations are considered later in the paper. Only 18 crops are presented in table 2. For a more detailed discussion of (a) price, yield, and gross income variance of 57 California crops, and (b) net income variability of selected cropping systems, see Carter, H. O. and Dean, G. W., *Relationships Between Income Stability and Income Levels for Principal California Crops and Cropping Systems*, Calif. Agr. Expt. Sta. Bul. (forthcoming).

[12] State annual average data for 1918–57 comprise the time series used. Data limitations are discussed in more detail later.

TABLE 1.—*Calculation of the variance of early fall lettuce yields in California, 1918–57, using the variate difference method*

Order of difference (1)	Unadjusted variance of the k th difference [1] (2)	Adjusted variance of the k th difference [2] (3)	Difference between adjusted variances (4)	H_{kN} (N = 40) [3] (5)	Standard error of difference between variances (6)	Standard error ratio [4] (7)
k	$V'_k = \dfrac{\Sigma \Delta_k^2}{N-k}$ (N=40)	$V_k = V'_k \div {}_{2k}C_k$	$V_k - V_{k+1}$	----------	${}^s V_k - V_{k+1} = V_k \div H_{kn}$ (N=40)	$R_k^\circ = \dfrac{V_k - V_{k+1}}{{}^s V_k - V_{k+1}}$
		Col. (2) $\div {}_{2k}C_k$	----------	----------	Col. (3) $\div$ Col. (5)	Col. (4) $\div$ Col. (6)
0--------	2,890.18	2,890.18	2,545.99	6.170	468.42	5.44
1--------	688.38	344.19	119.47	12.190	28.24	4.23
2--------	1,348.32	224.72	15.14	16.106	13.95	1.09
3--------	4,191.60	209.58	.99	18.860	11.11	.089
4--------	14,601.30	208.59	1.09	20.795	10.03	.109
5--------	52,290.00	207.50				

[1] Calculated as the sum of squares of the series of k th differences, divided by $N-k$.
[2] Further explanation available in Tintner, *op. cit.* p. 40–41.
[3] From table 20, Tintner, *op. cit.* p. 57–59.
[4] This test is based on the normal approximation.

most meaningful for comparisons between crops.[13] The empirical results emphasize the extremely wide range of variability resulting from the random elements associated with the diverse crops produced in California. Yield variability ranges from 2 percent for early fall tomatoes to 31 percent for olives; price variability ranges from 4 percent for wheat to 43 percent for early potatoes; gross income variability ranges from 7 percent for sugar beets to 36 percent for olives. High variability for many crops is even more striking in that the variability coefficients are based on the "random variance" which, in general, is considerably smaller than "total variance." Furthermore, use of aggregate data throughout (county and State data) probably causes an underestimate of variability facing individual farmers. (See appendix.) The results provide at least one reason for the development of a highly commercialized California agriculture. Many of the high-variability crops require highly technical specialized knowledge and, hence, tend to be produced by specialty growers. Only a producer with strong financial support can bear the risks associated with specialization in these high variability crops.

Consider the simplest case of combining two enterprises: Under the first method of diversification, total income variance (σ_T^2) is given by equation (2) in which σ_A^2 is the income variance

(2) $\qquad \sigma_T^2 = \sigma_A^2 + \sigma_B^2 + 2r_{AB}\sigma_A\sigma_B$

of enterprise A, σ_B^2 is the income variance of enterprise B and r_{AB} is the correlation between the incomes of enterprises A and B.[15] Under the second method of diversification, total income variance (σ_T^2) is given by equation (3) in which q is

(3) $\quad \sigma_T^2 = q^2\sigma_A^2 + (1-q)^2\sigma_B^2 + 2q(1-q)r_{AB}\sigma_A\sigma_B$

the proportion of resources devoted to A, $1-q$ is the proportion of resources devoted to B and the

[13] The question arises as to whether the random variance is homogeneous with respect to time. As a "rough" test, the random variances were computed for the two subperiods 1918–37 and 1938–57. Where no statistical difference was detected between these variances (as indicated by Bartlett's test of homogeneity of variance) the crop variance was based on the 1918–57 series; where the variance changed significantly over time, the variance based on the most recent 20-year subperiod was taken as the best estimate of future variance.

[15] Total income equals income from A plus income from B or: $T = A + B$

By definition: $\mathrm{Var}\,(T) = E[(A+B) - E(A+B)]^2$
$\qquad\qquad = \sigma_A^2 + \sigma_B^2 + 2r_{AB}\sigma_A\sigma_B$

For additional detail on derivation see, for example: Anderson, R. L., and Bancroft, T. A., *Statistical Theory in Research*, New York, New York: McGraw-Hill Book Co., Inc., 1952, p. 33.

TABLE 2.—*Relative variability of yield, price, and gross income of selected California crops*

Crop	Variability coefficients [1]		
	Yield	Price	Gross income
	Percent	*Percent*	*Percent*
Field crops:			
Alfalfa	3	11	15
Barley	5	10	8
Sugar beets	6	6	7
Potatoes, early	6	43	35
Wheat	7	4	8
Rice	10	10	10
Vegetables:			
Tomatoes, early fall	2	13	10
Tomatoes, processed	5	7	8
Onions, late summer	6	37	35
Lettuce, summer	9	18	31
Lettuce, winter	12	19	24
Cantaloups, spring	16	17	26
Fruits and nuts:			
Grapefruit	5	11	5
Peaches, clingstone	10	15	19
Grapes	10	31	29
Oranges, valencia	17	20	10
Almonds	19	21	17
Olives	31	27	36

[1] Variability coefficient $= \sqrt{\dfrac{\text{random variance}}{1953\text{–}1957 \text{ mean}}} \times 100$

other symbols are as defined above.[16] To estimate total variance in either case, estimates are needed of (1) the income variances of individual enterprises and (2) the correlation of incomes between enterprises.

As pointed out in the previous section, the variate difference method provides an estimate of the income variances of individual enterprises (that is, the variance of the random component of incomes). To be consistent with the concept of dealing with only the random element, the correlation coefficient should measure the association between the random components of the incomes of the two enterprises. Tintner [17] summarizes the logic and procedure for obtaining the correlation between the random elements of two time series. The approach is similar to that employed in obtaining the variance of a random component of

[16] Thus, $T = qA + (1-q)B$

$\text{Var }(T) = E\{qA + (1-q)B - E[qA + (1-q)B]\}^2$
$= q^2\sigma_A^2 + (1-q)^2\sigma_B^2 + 2q(1-q)r_{AB}\sigma_A\sigma_B$

Likewise, see Anderson, R. L., and Bancroft, T. A., *Ibid.*, p. 33.

[17] Tintner, *op. cit.*, p. 117–129.

a time series. Again, it is assumed that each series consists of a nonrandom element or mathematical expectation and a random element. The product moment and correlation coefficients between the two series are computed for each of the successive differences. This procedure is continued until the product moments (p_k) of the successive differences stabilize. If the nonrandom element has been eliminated in the $k_0{}^{th}$ difference, the following relationship holds:

$$(4) \qquad p_{k_o} = p_{k_{o+1}} = p_{k_{o+2}} = \ldots$$

If the difference between the product moments of two successive differences is smaller than about three times its standard error, it is reasonable to assume that the nonrandom elements have been eliminated, leaving an estimate of the product moment of the random elements.

Table 3 summarizes the computations involved in obtaining an estimate of the correlation coefficient between tomatoes and sugar beets in Yolo County, Calif. Using the standard error ratio criterion, the correlation between the first differences (0.05) is taken as the estimate of the correlation between the random elements of the two series. This correlation (0.05) is considerably less than that between the original two series (0.82). While both series exhibit strong upward trends, there is little correlation between the first differences.

Table 4 provides a summary of net income correlations (based on county yields and costs and State prices) as computed by the variate difference method between selected pairs of crops grown in several major farming areas of California. In general, the correlations obtained by the variate difference method are much lower than those between the original series. The actual net incomes of crops tend to be highly correlated because the major economic influences (inflation, price cycles, wars, level of technology, and so on), affect most enterprises similarly. The question is whether the correlation between (1) the original series or (2) the random elements is more meaningful for decision making. In the "no knowledge" case mentioned previously (p. 44), in which all deviations from the long-run mean are considered random or unpredictable, the relevant correlation is between the original series. As derivation of

TABLE 3.—*Calculation of the net income correlation between tomatoes and sugar beets in Yolo County, Calif., 1938–57, using the variate difference method*

Order of difference (1)	Adjusted product moment of the kth differences [1] (2)	Difference between adjusted product moments (3)	Standard error of difference between adjusted product moments [2] (4)	Standard error ratio [3] (5)	Random variance of sugar beet net income [4] (6)	Random variance of tomato net income [4] (7)	Correlation coefficient between random components (8)
k	$P_k = \dfrac{\Sigma^{(k)} XY}{(N-k)_{2k}C_k}$ $(N=20)$	$P_k - P_{k+1}$	$^sP_k - P_{k+1} = \dfrac{\sqrt{L_k}}{H_{kN}}$ $(N=20)$	$\hat{R}_k^\circ = \dfrac{P_k - P_{k+1}}{^sP_k - P_{k+1}}$	$V_k(X)$	$V_k(Y)$	$r_k = \dfrac{P_k}{\sqrt{V_k(X)V_k(Y)}}$
				Col. (3) ÷ Col. (4)	------	------	$\dfrac{\text{Col. (2)} \div}{\sqrt{\text{Col. (6)} \times \text{Col. (7)}}}$
0	663.31	658.23	173.53	3.79	1,354.76	481.56	0.82
1	5.08	7.20	9.24	.78	304.58	34.80	.05
2	−2.12	−3.04	6.25	−.49	287.80	25.87	−.02
3	.92	------	------	------	293.86	24.23	.01

[1] Calculated from the sum of cross products of the kth differences of the two series divided $(N-k)_{2k}C_k$. X and Y denote sugar beet and tomato net incomes per acre, respectively.
[2] See Tintner, *op. cit.*, p. 119–120 for definitions and explanations of L_k and H_{kN}.
[3] This test is based on the normal approximation.
[4] Computed as in table 1.

the variance of an enterprise combination requires estimates of both individual enterprise variances and the covariance between enterprises, mathematical consistency requires both variances and covariances to be based either on the original series or on the random elements of the series. That is, it is mathematically inconsistent to combine random variances and "actual" correlations, or vice versa. In practice, however, farmers are generally aware of long-run trends and hence are constantly revising plans in light of new technology and changing demands and price relationships. For this group, the correlations of the random elements seem to be more meaningful. At a given point in time, the farmer is aware of the general relative levels of income from various enterprises. What he desires is a measure of the relationship between random year-to-year changes in net income for various crops. For example, if two enterprises have a strong negative correlation between their random components, they might make an excellent diversification prospect for a particular year, even though the correlation between the original series is strongly positive.

Empirical Results

In practice, most cropping systems in California include more than two crops. Generaliz-

ing the two-crop case, it can be shown that the total variance equation for redistributing resources among n enterprises is:

$$(5) \qquad \sigma_T^2 = \sum_{i=1}^{n} q_i^2 \sigma_i^2 + 2 \sum_{\substack{i,j=1 \\ i>j}}^{n} q_i q_j r_{ij} \sigma_i \sigma_j$$

in which q_i $(i=1, \ldots n)$ is the proportion of resources devoted to the ith enterprise and $\Sigma q_i = 1$. Using this general formula, the total variances were computed for specific common cropping systems in six agricultural areas of California. The standard deviations ($\sqrt{\text{random variances}}$), 1953–57 mean net incomes and variability coefficients of these cropping systems are presented in table 5.[18] Again, county yields and costs and State prices were used. The results in table 5 indicate the wide range in levels and variabilities of incomes within and between areas of California. Extension personnel in farm management and agronomy have indicated considerable interest in utilizing such information on a "practical" level in the State.

---

[18] For a more complete presentation and interpretation of results, see Carter, H. O. and Dean, G. W., *op. cit.* This publication also develops further relationships between income levels and stability where the proportions of crops in the systems are allowed to vary.

TABLE 4.—*Net income correlation coefficients between selected crop combinations in major farming Areas of California* [1]

Crop combination	Area					
	Northern Sacramento Valley	Yolo	Fresno-Madera	Kings-Tulare Lake	Kern	Imperial Valley
Alfalfa–Cantaloups			0.02			
Alfalfa–Sugar Beets		−0.38	.09		0.45	0.02
Alfalfa–Cotton			−.29	−0.32	−.29	.13
Sugar Beets–Cantaloups			.14			
Cantaloups–Cotton			.14			
Cotton–Sugar Beets			−.10			.13
Barley–Alfalfa		.42		.37	.19	.63
Barley–Cotton				.28	.35	.44
Barley–Sugar Beets		−.04			.24	.14
Potatoes–Alfalfa					.59	
Potatoes–Cotton					.18	
Potatoes–Sugar Beets					−.17	
Potatoes–Barley					.23	
Rice–Wheat	0.40				.31	
Rice–Barley	.43					
Tomatoes (owner-operator)–Alfalfa		.60				
Tomatoes (leased)–Alfalfa		.56				
Tomatoes (owner-operator)–Barley		.16				
Tomatoes (leased)–Barley		.10				
Sugar Beets–Tomatoes (owner-operator)		.01				
Sugar Beets–Tomatoes (leased)		.05				

[1] Computed using the variate difference method. Thus, correlations are between the random components of the pairs of series.

Conclusions

Use of aggregate data force certain limitations on the interpretation of net income variability estimates for individual crops and cropping systems. (See Appendix.) Despite this limitation, variability estimates may provide a fairly good idea of the relative risk or uncertainty of crop alternatives. Finally, the variate difference method provides estimates of variability which, it is argued, have greater relevance to decision making by individual farmers than estimates derived by other commonly used methods.

Aggregation Problems

Data limitations ordinarily force use of aggregate data in estimating the variability of yields, prices, and incomes. Since interest is in variability to the individual farmer, this question arises: Do variability measures derived from aggregate data accurately reflect individual farm variability? Intuitively, some "averaging-out" of individual farm variability might be expected in the compilation of aggregate data. This problem is expected to be most severe with respect to yield variability. However, if interest is in random variability, it is also possible that individual farm yields are not ideal for analysis. For example, individual fluctuations in farm yields depend not only on random influences but also on conscious changes by the farmer in practices, levels of inputs, and so on, from year-to-year. Thus, it is not entirely clear what type of yield data are ideal for this type of investigation. Most researchers, however, probably would prefer variability measures based on individual farm data if available. Unfortunately, individual farm yield information of sufficient historical length for analysis is almost nonexistent. In the California study, for example, county yield and cost data and State prices were used as the only source available.

How serious, then, is the aggregation problem in estimating individual farm yield variability? It can be shown that, under certain assumptions, the bias depends on N (the number of farms comprising the aggregate) and ρ (the correlation between the random components of yields on these

TABLE 5.—*Net income variability comparisons between selected crop combinations in six areas of California (assuming a 560-acre farm)*

NORTHERN SACRAMENTO VALLEY

Crop combination [1]	Mean net income 1953-57 [2] (1)	Standard deviation [3] (2)	Variability coefficient [4] (3)
	Dollars	*Dollars*	*Percent*
1. R–R–R-B	39,154	10,696	27
2. R–R–R–W	39,715	10,304	26
3. R–R- R–F	34,530	9,990	29

YOLO COUNTY AREA

1. A–A–A–SB–T–SB	40,891	5,992	15
2. A–A–A–T–SB–B	36,982	6,530	18
3. A–A–A–T(L)–SB–B	29,731	4,872	17
4. A–A–A–SB–B–B	25,172	4,833	25

FRESNO-MADERA AREA

1. A–A–A–Ca	35,358	15,786	45
2. A–A–A–C–SB	36,943	7,795	21
3. A–A–A–C–C–SB	45,461	10,595	23
4. A–A–A–C–C–Ca	53,284	15,036	28
5. A–A–A–C–SB–Ca	43,764	12,320	28
6. A–A–A–SB	23,862	8,305	35
7. A–A–A–SB–Ca	33,897	13,132	39

KINGS-TULARE LAKE BASIN

1. A–A–A–C–B	36,361	7,006	19

KERN COUNTY AREA

1. A–A–A–P–P	54,415	50,523	93
2. A–A–A–C–C–C	68,186	14,011	21
3. A–A–A–SB–C–SB	42,381	11,256	27
4. A–A–A–C–C–P	65,492	22,008	35
5. A–A–A–C–B–P	50,137	21,840	44
6. A–A–A–C–P	57,585	25,833	45
7. A–A–A–SB–B–P	36,842	23,660	64

IMPERIAL VALLEY

1. A–A–A–C–C–SB	48,681	9,850	20
2. A–A–A–C–B–SB	35,638	7,683	22
3. A–A–A–C–C–C	55,647	12,549	23
4. A–A–A–C–C–B	42,207	9,862	23

[1] Assumes equal proportions of the 560-acre farm devoted to each crop in the crop combination. For example' R–R–R–B refers to 420 acres of rice and 140 acres of barley. Symbols are defined as: A=alfalfa, B=barley, C=cotton' Ca=Cantaloups, SB=sugar beets, R=rice, W=wheat, F=fallow, T=tomatoes (owner handles complete operation) T(L)=tomatoes (owner leases land in return for 17 percent of gross income), and P=potatoes.

[2] Net income is defined for this study as gross income minus variable costs only.

[3] Standard deviation is the square root of the estimated variance for the respective crop combinations as estimated by the variate difference method.

[4] The variability coefficient is the ratio of the standard deviation (col. 2) to the mean net income (col. 1), multiplied by 100.

farms). The yield (Y) on the ith farm in the year t might be written:

$$(6) \qquad Y_{it}=\pi_1+\beta_i t+e_{it}$$

in which $t=$time in years and $i=1, 2 \ldots N$. π and β are parameters and e is a random variable. The usual assumptions are made:

$$E(e_{it})=0 \text{ and } E(e_{it})^2=\sigma_i^2.$$

Averaging over all N farms gives:

$$(7) \qquad \overline{Y}_t=\overline{\pi}+\overline{\beta}t+\overline{e}_t.$$

It is assumed that $E(\overline{e}_t)=0$. However, the variance of $\overline{e}_t$ is:

$$(8) \qquad \text{Var}(\overline{e}_t)=\frac{1}{N^2}\left(\sum_{i=1}^{N}\sigma_i^2+2\sum_{\substack{i,j=1 \\ i>j}}^{N}\rho_{ij}\sigma_i\sigma_j\right)$$

Two assumptions might be made at this point. First is the assumption that all farms in the aggregate considered have a common variance $(\sigma_i^2=\sigma^2)$. Certainly a crop grown in widely differing geographic locations and climatic environments might not have a homogeneous variance among all locations. Thus, a single variance estimate based on State averages might not accurately represent the variance for any particular area. Use of aggregate data involving a smaller geographic area (for example, county data) makes this assumption more realistic. The second assumption is that the random components of yields between all possible pairs of farms are correlated to the same extent $(\rho_{ij}=\rho)$. Granting these two assumptions, equation (8) may be rewritten as:

$$(9) \qquad \text{Var}(\overline{e}_t)=\frac{1}{N^2}[N\sigma^2+N(N-1)\rho\sigma^2]$$

$$=\frac{\sigma^2}{N}[1+(N-1\rho)].$$

If $\rho=1$ (that is, perfect yield correlation of the random components between farms), then Var $(\overline{e}_t)=\sigma^2$. However, if $\rho<1$ (undoubtedly, a more realistic case), the estimated variance varies inversely with the number of farms making up the aggregate. With respect to factor and product prices, the correlation between farms is probably close to one; hence, the number of farms making up the aggregate price and cost series would affect the corresponding variance estimates very little. With respect to yields, however, this bias is probably not negligible and should be recognized in using aggregate data. Despite this limitation, a rough idea of *relative* variabilities of crops can probably be derived from aggregate data.

In addition to aggregation problems, an internal inconsistency exists when time series data are used to first estimate yield and price variability and then gross income and net income variability. For example, gross income per acre is defined as the product of price and yield per acre. Further, yield and price series are each assumed to consist of a systematic and random element as follows:

$$(10) \qquad Y_t=\pi+\beta t+e_t \text{ and}$$

$$(11) \qquad P_t=\theta+\delta t+\epsilon_t.$$

Y is yield per acre, π and β are parameters, t denotes "time," and e is a random variable. P is price per unit, θ and δ are parameters, and ϵ is a random variable. Accordingly, gross income per acre (GI) in year t would be written as:

$$(12) \quad GI_t=Y_t P_t=\theta\pi+\theta\beta t+\theta e_t+\delta\pi t$$

$$+\delta\beta t^2+\delta t e_t+\pi\epsilon_t+\beta t\epsilon_t+e_t\epsilon_t.$$

It is apparent from the product of error and non-error terms (e.g., $\beta t\epsilon_t$, $\delta t e_t$, and so on) that differencing or trend-fitting methods can never completely eliminate the systematic components from the error terms. Consequently, these procedures provide only an approximation to the random variance for gross income. The same problem applies to estimating net income variance.

Appendix

Variate Difference Method of Estimating Variability of Cropping Combinations

Diversification of combinations of enterprises often is recommended as a means of lessening income variability. Since diversification principles are discussed elsewhere,[14] only a brief review is necessary here to provide the foundation for an empirical application of the variate difference method. Basically, diversification can be accomplished either by (1) adding sufficient resources to include the new enterprise or enterprises without reducing the size of the present enterprises or by (2) redistributing a constant quantity of resources among more enterprises. These limitations are not associated exclusively

---

[14] See for example: Heady, E. O., *Economics of Agricultural Production and Resource Use.* (New York: Prentice Hall, 1952), Chapter 17, Heady, E. O., Kehrberg, E. W. and Jebe, E. H. *op. cit.*, pp. 661–667; and Carter, H. O. and Dean, G. W. *op. cit.*, pp. 33–39.

with the variate difference method but rather with all methods of trend removal. However, one empirical problem relates particularly to the variate difference method: It is sometimes difficult to select the appropriate difference at which the variance of a series or the product moment of two series is stabilized. The standard error ratio criterion is not strictly applicable to "short" time series (small samples); hence, the choice of the random variance or product moment may involve some judgment on the part of the investigator.