Rob King

RISK ANALYSIS FOR AGRICULTURAL PRODUCTION FIRMS: CONCEPTS, INFORMATION REQUIREMENTS AND POLICY ISSUES

Proceedings of a Seminar Sponsored by
Southern Regional Project S-180
"An Economic Analysis of Risk Management
Strategies for Agricultural Production Firms"
New Orleans, Louisiana
March 25-28, 1984

Department of Agricultural Economics
Agricultural Experiment Station
College of Agriculture
University of Illinois at Urbana-Champaign

July, 1984

AE-4574

MEASURES OF DIVERSITY: A COMPARISON

Ray D. Nelson

Introduction

Both positive and normative studies of growth and investment decisions under uncertainty suffer from the chronic problem of how to measure diversification. The challenge arises not from a dearth of contending measures but rather from a plethora of candidates which often overwhelms research resources. The multitude of possible measures arises from two principal lines of reasoning. First, concentration based indices usually utilized in industrial organization to measure market power provide one set of viable measures. Second, variance based measures from the more traditional finance and risk management literatures comprise a another set of alternatives.

The concentration indices and variance estimators considered in the present study certainly do not exhaust the diversification measurement possibilities. Since no consensus establishes the superiority of a single measure, researchers must grapple with the dilemma of which statistic to use. Employing all, or even a large subset, of the possibilities yields robust results, yet quickly renders intractable research. Unless known similarities allow selection of representative measures from homogenous groupings, using a single or few carelessly and capriciously chosen measures may compromise both the validity and reliability of a study.

The present paper surveys a variety of diversification indices and then examines their validity and redundancy by asking the following three questions:

1) How much do each of the measures of diversity have in common?
2) What redundancy exists among the measures of diversity?
3) Do the measures of diversity validly measure some underlying phenomenon associated with risk management?

Previous work done in industrial organization, finance, psychology, and multivariate statistical analysis all contribute to the answers to these questions. Since this paper draws on topics expounded in such a wide variety of disciplines, many of the well-known concepts of concentration, risk measurement, and multivariate techniques are briefly described.

Ray D. Nelson is an assistant professor in the Department of Agricultural Economics at the University of California, Davis.

Units of Measure

A discussion of measures of diversification first requires a clear understanding of the importance of units of measure. The choice of units means the selection of resources, assets, revenues, or profits to represent the distribution of firm effort to all possible activities. Each of the variables usable to calculate proportions of effort has both advantages and disadvantages. The availability of planted acreage data makes land an especially attractive farm resource or asset to use as a unit of measure. Land does not, however, account for different crop values. Proportions based on revenue incorporate crop values but ignore the possible association between large revenues and commensurately high input costs. Inaccurate farm records and inconsistent accounting methods cause imprecision in proportion calculations based on revenues and profits.

Pope and Prescott illustrate the possible differences in proportions which can arise from using different units of measure. Table 1 presents a similar example. The data represent a farm allocating equal acreage to producing rice, corn, and tomatoes.

Despite the recognition of the importance of the units of measure in determining farm diversity, unavailable or inaccurate data often limit researchers to using only planted acreage data as the sole indicator of firm effort devoted to a given activity. Even with the existence of a good time series of revenue per acre, the lack of corresponding estimates of the average costs of production precludes analysis based on profits per acre.

The focus in the present study on the similarities of candidate indices of diversity makes measurements based on revenues and profits as well as acreage all the more important. For the empirical testing which utilizes Sacramento-California data, annual County Agricultural Commissioner Reports provide the price and yield information needed to calculate the revenue per acre for each cropping activity. The estimation of costs per acre required to determine annual profits depends on a longitudinal data set composed of farm budgets calculated by county farm advisers and University of California extension specialists. Nonsystematic accounting and inconsistent economic methodologies contribute substantial random noise to these observations.

The estimation of average variable costs per acre assumes that production expenses change over time because of variation in the overall price level, producer input prices, and trends in production techniques. More formally, the regression model holds that the real average costs of production depend on a dummy variable representing membership of a given observation in the Sacramento Valley, a linear time trend, and the quotient of the producer input index and the CPI. Three other dummy variables represent special conditions in the production of drybeans and tomatoes. Table 2 summarizes the regression results.

TABLE 1

Proportions of Farm Firm Effort

	Acreage (acres)	Revenue (Dollars)	Profit (Dollars)
Crops			
Rice	100	30,000	20,000
Corn	100	10,000	8,000
Tomatoes	100	100,000	30,000
Total	300	140,000	58,000
Proportions			
Rice	.33	.214	.345
Corn	.33	.071	.138
Tomatoes	.33	.714	.517

TABLE 2

Estimated Average Variable Cost of Production Equations*

Crop	Intercept	Sacramento Intercept	Time	Producer Price Index	Drybean Intercept	Tomato Intercept	Tomato Slope
Barley	-1561.30	-42.33	.66	413.95			
R^2 = .30	(3370.49)	(18.29)	(1.76)	(161.11)			
Drybeans	-23274.20	22.46	11.93	139.75	-51.14		
R^2 = .31	(6304.54)	(55.87)	(3.28)	(327.94)	(55.87)		
Corn	-10294.50	-5.58	5.17	386.47			
R^2 = .44	(3345.55)	(17.95)	(1.74)	(151.62)			
Alfalfa	-5041.6	-68.82	2.80	-184.94			
R^2 = .20	(4510.74)	(38.68)	(2.32)	(222.75)			
Pasture	124.80	-45.46	-.13	257.31			
R^2 = .31	(2260.35)	(19.76)	(1.17)	(100.52)			
Rice	-9871.9	-56.89	5.19	10.34			
R^2 = .37	(4743.01)	(26.78)	(2.46)	(223.32)			
Safflower	-3233.38	-107.35	1.67	149.64			
R^2 = .53	(4145.13)	(22.24)	(2.14)	(161.92)			
Sorghum	-5006.81	-23.06	2.55	175.68			
R^2 = .44	(1649.37)	(10.25)	(.85)	(72.80)			
Sugar Beets	-14942.50	-107.11	7.64	490.23			
R^2 = .35	(4959.39)	(32.35)	(2.59)	(271.21)			
Wheat	-1561.30	-42.33	.66	413.95			
R^2 = .30	(3370.49)	(18.29)	(1.76)	(161.11)			
Almonds	-13381.50	-166.208	7.00	489.68			
R^2 = .38	(8445.80)	(55.78)	(4.35)	(431.43)			
Tomatoes	-251204.00	-186.91	128.61	-159.11		237264.20	-120.896
R^2 = .25	(96342.44)	(76.03)	(49.04)	(729.1)		(102680.60)	(52.28)

*Values in parentheses below estimates represent standard deviations.

Industrial Organization Measures

Concentration indices meter levels of naive diversification. Although primarily developed and used by industrial organization economists to determine market concentration, the many indices sometimes find use in gauging a firm's diversity. Pope and Prescott constitute one such application.

Industrial organization specialists have meticulously explored industrial concentration and firm specialization. Luckily, Curry and George summarize this immense literature in their industrial concentration survey. Since the market shares of firms within an industry determine concentration and the relative allocation of resources within a firm determines specialization, both topics utilize a common set of indices. Although these indices share a dependence on size distributions, their differences arise from the preoccupation with markets in concentration and the emphasis on the firm in the case of specialization.

Concentration Measurement

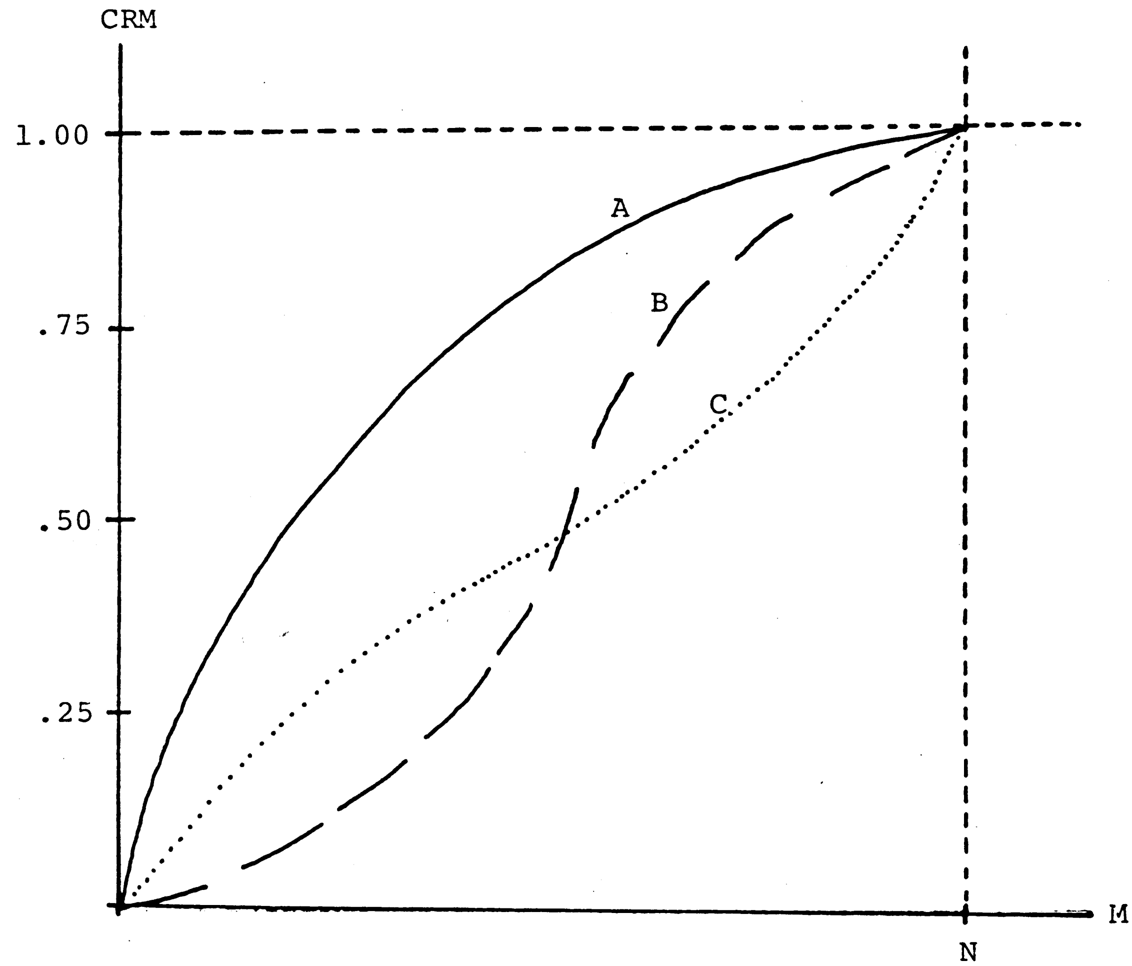
Concentration measures attempt to reduce all the data in a size distribution into a single statistic. Since no such statistic perfectly summarizes the information in a distribution, abundant alternatives which capture different dimensions of concentration provide researchers a choice of a wide range of measurement tools.

In the present study, the proportions X_i of resources, revenues, or profits dedicated to or derived from the ith largest production activity form the size distribution. Since the individual proportions often fail to contribute the perspective available in an entire cumulative density, defining the cumulative share of the largest M activities i = 1,. . ., M as $CRM = \sum_{i=1}^{M} X_i$ allows a more comprehensive comparison of specialization of firms which have the same production possibilities. The graphs in Figure 1 illustrate such comparisons for three firms with N possible activities. Because at any $M<N$ the cumulative share of activities one through M for A always exceeds that of B and C, the specialization of the first surpasses that of the other two. Since the cumulative curves for B and C intersect, however, cumulative concentrations do not establish dominant specialization.

The widespread practice of arbitrarily choosing a value of M for the concentration ratio reduces the distribution to a single statistic. Such a statistic demonstrates the convenience and danger of single statistic summaries of multiple dimensions. A single value of M facilitates concentration ratio calculation and ranks the specialization of each firm. Researchers utilizing only traditional values of M encounter no dissonance unless experimenting with alternative values of M alters previous orderings. Despite its shortcomings, the concentration ratio remains a very popular measure.

Figure 1

Cumulative Size Distribution



Concentration Indices

The popularity of the concentration ratio cannot be attributed to a lack of alternatives. Diligent efforts to produce indices congruent with economic theory and consistent with sets of desirable properties[1] have produced many possible indices. Marfels suggests that these indices are all functions of the proportions X_i and number of activities N. Each index simply applies different weights to the proportions. The weights themselves often depend on the values of X_i. Since the number of activities can affect the size of the X_i's, Marfels shows that increasing N can directly and indirectly change the measure of concentration. Therefore, comparisons of concentration must account for the possibility of an unequal number of activities.

Aaronnovitch and Sawyer, Curry and George, Hannah and Kay, and Marfels all summarize the prominent measures of concentration. Table 3 lists a select subset of these measures along with their formulas, extreme values, and relationship to diversification. The formulas and extreme values in Table 3 illustrate two important points. First, the weights given to the proportions differentiate the measures from each other. The Herfindahl index, for example, weights each proportion by the proportion itself. This tends to allocate heavy weight to large activities. In contrast, the smaller activities receive heavier weights ($-\ln X_i$) in the entropy index. The extreme values in Table 3 demonstrate the second point which is that every index has either a maximum or minimum determined as a function of the number of possible activities. Hence, caution should temper comparison of specialization of firms with different numbers of possible activities.

Markowitz Diversification

Concentration indices indicate naive assessments of risk management due to combinations of alternative activities. Because activities most likely fail to satisfy the Samuelson assumption of identical, independent distributions, these naive measures may not indicate the degree to which diversification limits risk. Markowitz or variance based portfolio measures may, however, reflect the degree of risk management effectiveness.

Like many other approaches, Markowitz employs variance to assess the riskiness of combinations or portfolios of activities. Determining the level of portfolio risk requires estimates of the variances and covariances of all activities as shown in the following formula:

$$\sigma_p^2 = \sum_i \sum_j X_i X_j \sigma_{ij}$$

where σ_p^2 is the variance of the portfolio,

TABLE 3

Indices of Diversity

Index Name	Formula	Min	Max	Correlation with Diversification
Concerntration Ratio $M = 1,\ldots,4$ (CRM)	$CRM = \sum_{i=1}^{M} X_i$	$\frac{1}{N}$	1	Negative
Herfindahl (H)	$H = \sum_{i=1}^{N} X_i^2$	$\frac{1}{N}$	1	Negative
Hall-Tideman (HT)	$HT = \frac{1}{2 \sum_{i=1}^{N} i\, X_i - 1}$	$\frac{1}{N}$	1	Negative
Entropy (E)	$E = -\sum_{i=1}^{N} X_i \ln X_i$	0	$\ln N$	Positive
Comprehensive Concentration (CCI)	$CCI = X_1 + \sum_{j=2}^{N} X_j\, [1-(1-X_j)]$	1	$\frac{(N-1)(2N-1)}{N^2}$	Positive

[a]Because some of the measures are not symmetric, the activities are ranked in descending order with the ith activity receiving the ith rank.

X_i and X_j are once again the proportions of the ith and jth activities, and

σ_{ij} is the covariance between the rates of return of the ith and jth activities.

Estimation of the variances and covariances from _ex post_ times series data supposes that the observations result from a stochastic process. The stationarity of the process determines the correct estimation technique. Since stationarity requires that the probability distribution from which observations are drawn be time invariant, a stationary stochastic process cannot exhibit a trend nor include periods of varying volatility. For these reasons, many economic times series do not qualify as stationary stochastic processes. Many can nonetheless be transformed so that they become stationary or approximately stationary.

Many different alternative procedures provide estimates of the variance and covariance parameters. The following three find direct application in the present analysis:

1) Simple _ex post_ estimation,
2) Variate difference method, and
3) Single Index Model

Simple Ex Post Estimation

Well established trends in prices and technology suggest a need for transformations of revenue and profit time series in order to achieve stationarity. The effect of inflation must obviously be eliminated through deflating prices in order to account for the upward trends in nominal revenues and profits.

Assuming the stationarity of real revenues and profits facilitates the calculation requisite for measuring effective Markowitz diversification. Under this assumption, simple, well-known formulas unbiasedly estimate long-run variances and covariances. Since the stationarity assumption of the times series is questionable, it should be investigated. Such an investigation utilizes parts of the variate difference procedure.

Variate Difference Method

Tintner details the variate difference method and Dean and Carter demonstrate its applicability to agriculture. Rather than taking the more traditional approach of decomposing a time series into a trend, cycle, season, and random error, Tintner suggests a two component alternative. Trends, cycles, and season together form the first component which captures the permanent variation in the time series. When graphed with respect to time, this mathematically expected variation forms a smooth, continuous curve whose consecutive elements exhibit strong correlation. The random or second component encapsulates from finding the simple average in each year of all individual indices representing each possible activity in the region. Regressing the individual crop revenue and profit indices on the average composite gives the needed beta estimates.

Two other index constructions assume that state or national revenue or income figures may better represent systematic risk. The calculation of these four additional indices utilizes the same process employed to generate the single activity indices. The beta estimation procedure then requires regressing the individual crop indices on those constructed from California and United States farm revenue and income data. Table 4 reports the results of regressing each crop index on the six possible measures of systematic risk.

Since estimation of the single index parameters usually utilizes time series data, stationarity can still compromise the value of the model. This questionable stationarity has, in fact, inspired numerous investigations. Beta adjustment procedures which make better _ex ante_ predictions of variances and covariances have resulted from these studies.

Common Factors and Similarities

The possible combinations between the acreage, revenue, or profit units of measure and the industrial organization or Markowitz originated indices offers researchers thirty-four different alternatives to measuring diversity. Applying the eight formulas enumerated in Table 3 to each of the three possible units of measure gives twenty-four industrial organization indices. Using the simple _ex post_ or variate difference methodogies with revenue or profit data yields four portfolio variance measures. Weighting the six different sets of beta estimates with the appropriate acreage proportions contributes the final six alternatives.

Multivariate procedures provide the methodology needed to compare and evaluate these diversity measures. Application of factor analysis, a technique originally developed to distill intelligence quotients from exam data, to thirty-four different diversification indices estimated for a random sample of 211 Sacramento Valley farms tests for a common, underlying diversity factor. Cluster analysis, a procedure with multidisciplinary applications, categorizes the measures according to similarity. Although the possible redundancy of employing both techniques to the same data set may be an issue with some, using factor and cluster analysis together adds confidence to the conclusions reached.

A Diversity Factor

De Leeuw's recent highlighting of psychometric methods relevant to econometric analysis supports the application of factor analysis to the present problem. Spearman originally developed factor analysis to all nonpermanent and unexpected variation and manifests no autocorrelation.

The variate difference method prescribes successive differencing of the ordered observations to eliminate the permanent component. Since the random errors are uneffected by order, the residuals of the process represent the stochastic component. Applying unbiased estimators to these residuals yields the needed variances and covariances. When successive differencing steps yield stable estimates, the variate difference method is complete. The estimates in the present study stabilize after the first differencing operation.

Single Index Model

Econometric models condition expectations on the values of related variables. One group of very simple econometric formulations, the index models proposed by Sharpe, not only facilitate computation but also decompose risk into two components. The decompostition of risk designates the explained variation as systematic and the residual as unsystematic. The index model has the additional twin advantages of significantly reducing the calculations needed to estimate the variances and covariances and greatly simplifying the computation of optimal portfolios.

The single index model seems applicable to agriculture since common economic, agronomic, technological, and meteorological forces engender the observed high positive correlations among production-marketing activities. A general index representing a portfolio of all activities captures the variation due to common trends, cycles, and seasons. Regressing the returns to a single activity on the index identifies the systematic risk component. If r_i and r_I, respectively, represent the returns on the ith activity and index, then the regression equation becomes:

$$r_{it} = \alpha_i + \beta_i \, r_{It} + \varepsilon_{it}$$

The portfolio beta results from finding the weighted sum of the betas corresponding to each production activity. The acreage proportions serve as the weighting factors as reflected by the following formula:

$$\beta = \sum_{i=1}^{N} X_i \cdot \beta_i$$

As in the case of the application of the single index model to securities, multiple possible indices arise to represent the systematic risk. Latane, Tutle, and Young detail buy and hold construction methodology in their discussion of the choice of an optimal index. Implementation of the buy and hold process requires that production activity revenues and profits at each point in time be considered relative to their values in a given base year. The annual values for indices representing systematic risk for a given geographic area result

TABLE 4

Estimated Beta Coefficients

	Revenue			Income		
	Glenn-Colusa	California	U.S.	Glenn-Colusa	California	U.S.
Barley	0.88	1.11	1.12	1.28	1.65	2.82
Drybeans	1.02	1.08	1.21	0.94	1.77	1.87
Corn	0.85	0.79	1.04	0.54	1.36	1.40
Alfalfa	0.89	0.84	0.97	0.93	1.97	2.25
Pasture	1.55	2.28	2.28	2.29	5.66	1.67
Rice	1.46	2.25	2.29	1.19	3.78	1.11
Safflower	0.01	-0.04	-0.08	-0.28	-1.44	-0.49
Sorghum	0.98	1.20	1.50	0.51	1.34	1.91
Sugar Beets	1.56	1.60	1.51	2.27	3.67	0.98
Wheat	0.01	0.26	0.45	-0.13	0.36	1.40
Almonds	0.38	0.22	0.17	0.26	0.49	0.44
Tomatoes	2.41	2.54	2.46	2.20	4.13	1.32

condense myriads of examination scores into measures of intelligence. In general, this approach searches for common structure among sets of observed and measured variables. Rather than partitioning the measured variables and conditioning them on one another through regression, this analysis supposes that unobserved factors determine the value of each variable. The method assumes additivity of factors as shown in the following equation:[2]

$$X_{ik} = \lambda_{i1} \cdot f_{1k} + \lambda_{i2} \cdot f_{2k} + \ldots + \lambda_{im} \cdot f_{mk} + e_{ik}$$

where X_{ik} = kth observed value of the ith variable;

f_{jk} = value of the jth factor corresponding to the kth observation (known as factor scores);

λ_{ij} = relationship of the ith variable to the jth factor (known as factor loadings);

e_{ik} = random noise associated with the kth observation and the ith variable.

Factor analysis estimates the factor loadings and scores in a joint attempt to:

1) simplify the analysis by reducing the number of variables, and
2) discover the underlying structure of the data.

Measuring diversity involves many of the same problems inherent in measuring intelligence. Just as factor analysis has played a major role in the study of intelligence, its use can also shed light on diversification. Analysis results showing only one significant factor with estimated loadings of the anticipated signs for each proposed index would please researchers proposing to measure diversity. This single factor could then assume a designation attributable to risk management and researchers could feel heartened about the validity of their measures.

The summary statistics shown in Table 5 give the results of a factor analysis of the thirty-four possible diversification measures calculated for the sample of Sacramento Valley farms. The number of factors estimated utilizes the rule of thumb which excludes all factors with eigenvalues less than one.

Although the large variance explained by the first factor suggests that all the indices meter a common underlying phenomenon, significantly large loadings on the second, third, and fourth factors indicate that diversification indices include more than just risk management. In fact, because the variance based measures load so heavily on the second, third, and fourth factors, the first factor cannot justifiably receive the risk management designation. This result raises serious questions with regard to the validity of diversification indices.

Table 5

FACTOR ANALYSIS, PRINCIPAL COMPONENTS, NO ROTATIONS.

INITIAL FACTOR METHOD: PRINCIPAL COMPONENTS

	FACTOR PATTERN					
	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	
CR1A	0.91593	0.05958	0.02378	-0.21623	-0.22663	CR1 ACREAGE
CR2A	0.78741	0.34357	-0.00721	0.19376	0.18237	CR2 ACREAGE
CR3A	0.57250	0.52316	0.03235	0.51975	0.14412	CR3 ACREAGE
CR4A	0.30535	0.50893	0.09493	0.68362	-0.39070	CR4 ACREAGE
HA	0.91447	0.03938	0.01926	-0.21737	-0.23373	HERFINDAHL ACREAGE
HTA	0.91767	0.05049	0.02081	-0.21204	-0.22339	HALL-TIDEMAN ACREAGE
EA	-0.93290	-0.09491	-0.01707	0.14559	0.17884	ENTROPY ACREAGE
CCIA	-0.94337	-0.10236	-0.01586	0.14661	0.17130	CONCENTRATION ACREAGE
CR1R	0.94077	0.03790	-0.01306	-0.23500	-0.05673	CR1 REVENUE
CR2R	0.79733	0.31020	-0.04881	0.15176	0.39412	CR2 REVENUE
CR3R	0.61088	0.44636	-0.01981	0.38114	0.45923	CR3 REVENUE
CR4R	0.30158	0.50873	0.09539	0.68384	-0.39370	CR4 REVENUE
HR	0.95915	0.03209	-0.00578	-0.22214	-0.11718	HERFINDAHL REVENUE
HTR	0.96787	0.06209	-0.00359	-0.18851	-0.09882	HALL-TIDEMAN REVENUE
ER	-0.97584	-0.07706	0.00498	0.16573	0.08170	ENTROPY REVENUE
CCIR	-0.96936	-0.06149	0.01383	0.19231	0.04825	CONCENTRATION REVENUE
CR1P	0.91806	0.00877	-0.07532	-0.23842	0.05676	CR1 PROFIT
CR2P	0.74204	0.22155	-0.07772	0.04907	0.51570	CR2 PROFIT
CR3P	0.54825	0.33323	-0.05299	0.23043	0.62000	CR3 PROFIT
CR4P	0.30194	0.50877	0.09535	0.68385	-0.39344	CR4 PROFIT
HP	0.94706	-0.00642	-0.06007	-0.22755	-0.02123	HERFINDAHL PROFIT
HTP	0.95855	0.02118	-0.05627	-0.20102	-0.01481	HALL-TIDEMAN PROFIT
EP	-0.96928	-0.02907	0.05033	0.18974	-0.00112	ENTROPY PROFIT
CCIP	-0.95050	-0.01437	0.06573	0.21049	-0.04717	CONCENTRATION PROFIT
VCOVR	0.39400	-0.41941	0.79968	0.15349	0.05863	VARCOVAR REVENUE
VCOVP	0.40204	-0.41554	0.79907	0.14791	0.05517	VARCOVAR PROFIT
VDIFFR	0.38283	-0.43164	0.79278	0.16543	0.05963	VARIATE DIFFERENCE REVENUE
VDIFFP	0.38969	-0.42694	0.79439	0.15914	0.05577	VARIATE DIFFERENCE PROFIT
GCR	-0.48230	0.73451	0.16417	-0.39656	-0.02179	GLEN COLUSA REVENUE
GCP	-0.56840	0.67433	0.28402	-0.32917	-0.00368	GLEN COLUSA PROFIT
CAR	-0.50666	0.68032	0.37725	-0.36085	0.00345	CALIFORNIA REVENUE
CAP	-0.58859	0.67155	0.22426	-0.35008	-0.02023	CALIFORNIA PROFIT
USR	-0.53057	0.62175	0.43857	-0.33322	0.00760	UNITED STATES REVENUE
USP	-0.22666	0.33424	0.76400	-0.28165	0.07178	UNITED STATES PROFIT

VARIANCE EXPLAINED BY EACH FACTOR

FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5
18.603006	4.792279	3.675413	3.356526	1.811365

FINAL COMMUNALITY ESTIMATES: TOTAL = 32.238589

CR1A	CR2A	CR3A	CR4A	HA	HTA	EA	CCIA	CR1R	CR2R	CR3R	CR4R
0.941161	0.808904	0.893413	0.981237	0.940049	0.939959	0.932787	0.951506	0.945089	0.912704	0.928959	0.981496
HR	**HTR**	**ER**	**CCIR**	**CR1P**	**CR2P**	**CR3P**	**CR4P**	**HP**	**HTP**	**EP**	**CCIP**
0.984107	0.985939	0.992365	0.982939	0.908658	0.874103	0.851933	0.981557	0.952801	0.963059	0.978890	0.954513

The factor analysis suggests two other noteworthy empirical results. First, comparison of the acreage, revenue, and profit loadings for each different type of concentration based index reveals very strong similarities. This indicates less significance of the units of measure than previously brought into question by the example shown in Table 1. The second empirical observation comprises a very curious result. The variance and beta based measures load with opposite signs on the first, second, and fourth factors. Problems with the single index representation of systematic risk or beta estimation methodology possibly account for this anomaly.

Clusters of Similar Measures

Although few question the appropriateness of dissecting individual datum in search of underlying structure by using factor analysis, some discourage its use to reduce variable numbers to a manageable size.[3] Cluster analysis provides an alternative method which associates groups of similar whole objects (variables or observations). Its application, therefore, seems suitable for collapsing the multitude of diversification measures into a simple few.

Cluster analysis actually encompasses a variety of aggregating and disaggregating techniques which group entities according to similarity. The method strives to form clusters which are heterogeneous between but homogenous within. One subset of techniques called hierarchial agglomerative illustrates the general cluster analysis concept. This technique begins by treating each entity as a separate cluster and then sequentially forms groups until a single cluster includes all variables or observations. For example, in the first step using the nearest neighbor algorithm, the pair of closest entities form the first cluster. In the second step, the algorithm searches for the next closest entities. If the pair includes a member of an existing cluster, the outside entity join the existing cluster. The pair become a new group when neither are part of an existing cluster. When both elements of the pair already have membership in an existing cluster, the two clusters merge to form a single group. The process proceeds until all clusters and entities coalesce into one set.

The dendrogram shown in Figure 2 graphically summarizes the results of a cluster analysis of the thirty-four different diversification measures. Jackson details the general construction and interpretation of such diagrams. In the dendrogram of the present study, the variables named at the top of the diagram correspond to the same definitions given in Table 5. The vertical axis communicates the level of similarity among variables.

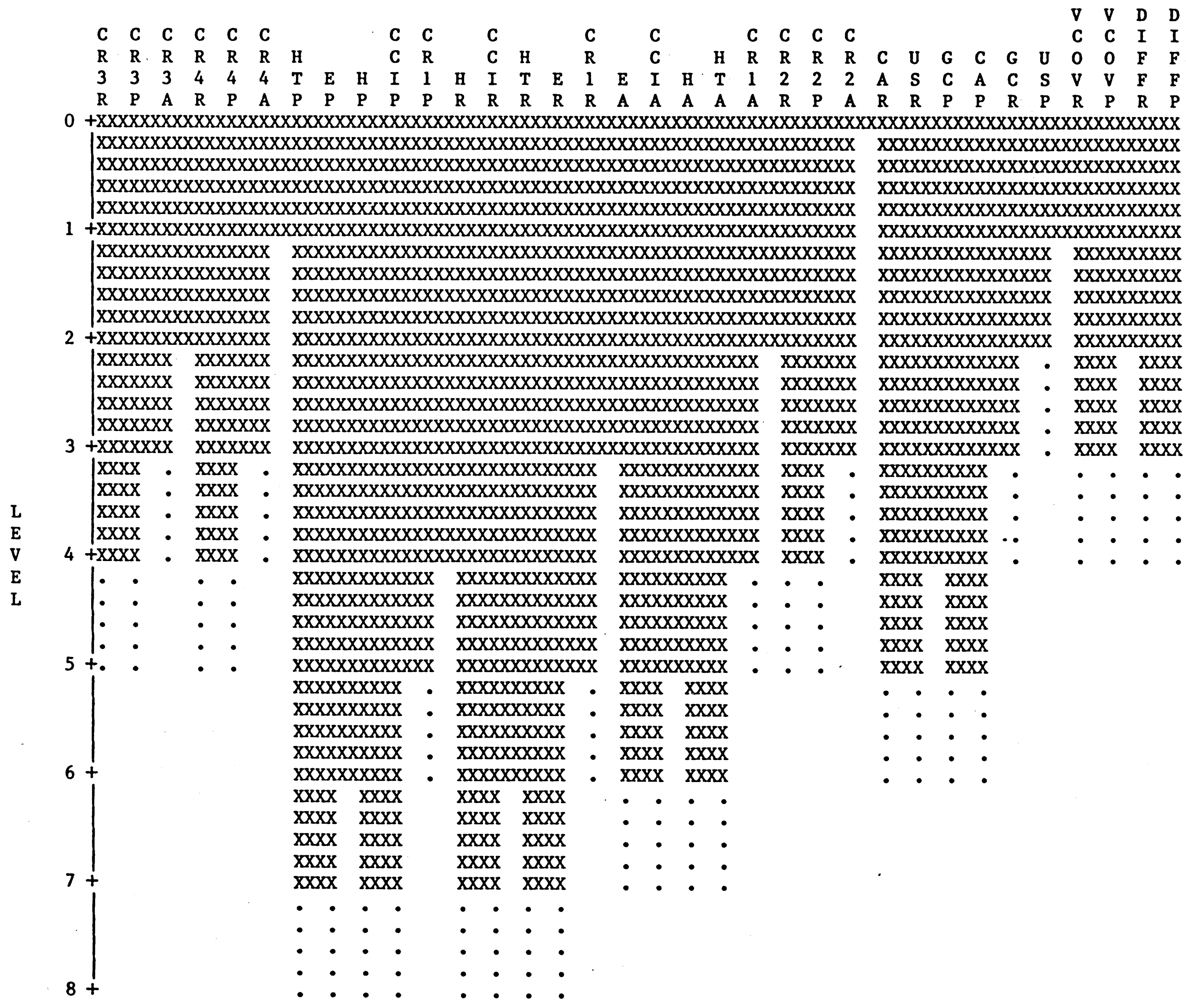
Interpreting the dendrogram from the bottom towards the top shows that the cluster analysis first pairs Hall-Tideman with entropy and comprehensive concentration with Herfindahl. This occurs for the indices based on both profits and revenues. The 'XXXX' notation at level seven of similarity denotes the formation of these first four

Figure 2

Dendrogram of Concentration Measures

OBLIQUE PRINCIPAL COMPONENT CLUSTERING

NAME OF VARIABLE OR CLUSTER



groups. At level six of similarity, the dendrogram reveals the cluster analysis merges Hall-Tideman, entropy, Herfindahl, and comprehensive concentration into one group. Once again this result occurs separately for both the revenue and profit based measures. At this same sixth level of similarity, the clustering process pairs the entropy with comprehensive concentration and Herfindahl with Hall-Tideman for the indices derived from acreage data. The grouping process continues until all measures find membership in a single set at the zero level of similarity at the top of the dendrogram.

Scrutinizing the dendrogram at four of the different levels of similarity yields some noteworthy information. The evidence at level five shows that many of the industrial organization based indices find membership in a given set based on their derivation from either profit, revenue, or acreage units of measure. This suggests that units of measure cause more differences among these indices than do the index types. This similarity among the industrial organization types indicates their redundancy.

The level four groupings reveal an important fact about the methods used to estimate variances and covariances. The far-right side of the dendrogram indicates two sets whose membership depends on method of calculation rather than the units of measure of the data utilized for estimation. This means that the choice between simple <u>ex post</u> and variate difference estimation methods is more important than the selection of revenue or profit as units of measure.

When the cluster analysis proceeds to level two, all measures find membership in one of four groups. These groups consist of:

1) those concentrations ratios greater than two,
2) all other industrial organization contributed measures,
3) indices defined as the portfolio's beta value, and
4) statistics based on estimated variances and covariances.

At level one, all industrial organization indices merge into a single group and all Markowitz measures of diversification join to form one set. This last result suggests that the methodological differences between the two approaches to measuring diversity translate into empirical differences as well.

<u>Conclusions</u>

The present research suggests that although all measures of diversification do seem to measure an underlying phenomenon, additional factors such as crop rotations, heterogenous land, managerial skills, and personal preferences also may influence the allocation of farm firm effort to each of the possible cropping activities. The strength of the loadings on four factors brings into question whether either industrial organization or Markowitz based indices validly meter risk management.

Despite the numerous candidates for diversity measures, the questionable validity of the current indices highlights the need for new alternatives.

Until such measures are forthcoming, researchers studying farm diversity and its relationship to other factors may want to utilize both industrial organization and Markowitz based indices. Because of the high degree of redundancy among the industrial organization class of indices, however, a representative such as the popular entropy index sufficiently typifies this group. In order to avoid the tedious costs of production estimations needed to calculate profits, revenue rates as the practical choice for units of measure. The early association of revenue and profit indices in the cluster analysis and their similar loadings in the factor analysis support this simplification. Given the availability of acreage data and the need to use acreage proportions in the estimation of portfolio variances, calculating an industrial organization type measure based on acreage proportions adds robustness with little additional effort. Among the Markowitz measures, the appropriateness of the variate difference method exceeds all others because of the nonstationarity of the data. Until perfection of the beta estimation procedure, the concept of the beta of the farm portfolio should only find use with extreme caution.

Footnotes

[1]Hause exemplifies attempts to fit concentration measures into economic theory. Hall and Tideman, as well as, Hannah and Kay specify the set of desirable properties that indices should satisfy.

[2]Extensive exposition of factor analysis methods has no place in the present paper. Morrison offers a comprehensive statistical treatment of factor analysis. Jackson gives an excellent description of its practical application.

[3]Jackson discusses the appropriate uses and misuses of both factor and cluster analyses.

References

Aaronovitch, S. and M. C. Sawyer. Big Business: Theoretical and Empirical Aspects of Concentration and Mergers in the United Kingdom. MacMillan Press: London, 1975.

Curry, B. and K. D. George. "Industrial Concentration: A Survey." The Journal of Industrial Economics, Vol. 31, #3 (March 1983): 203-255.

Dean, G. W. and H. O. Carter. "Measure of Enterprise Variability by the Variate Difference Method." Agricultural Economics Research, Vol. 7, #2 (April 1960): 43-52.

De Leew, J. "Editor's Introduction." Journal of Econometrics. Vol. 22 (1983): 1-12.

Hall, M. and N. Tideman. "Measures of Concentration." Journal of the American Statistical Association, Vol. 62 (March 1967): 162-68.

Hannah, L. and J. A. Kay. Concentration in Modern Industry: Theory, Measurement, and the U.K. Experience. MacMillan Press, London: 1977.

Hause, J. C. "The Measurement of Concentrated Industrial Structure and Size Distribution of Firms." Annals of Economic and Social Measurement, Vol. 6, #1 (1977): 73-107.

Jackson, B. B. Multivariate Data Analysis: An Introduction. Homewood, Illinois: Irwin, 1983.

Latane, H., D. Tutle, and A. Young. "How to Choose a Market Index." Financial Analysts Journal, Vol. 27, #4 (September-October 1971): 75-85.

Marfels, C. "The Consistency of Concentration Measures: A Mathematical Evaluation." Zeitschrift Fur Die Gesamte Staatswissenschaft. Quni, (1972): 196-215.

Markowitz, H. M. Portfolio Selection: Efficient Diversification of Investments. New York: John Wiley & Sons, 1959.

Morrison, D. F. Multivariate Statistical Methods. McGraw-Hill: New York, 1976.

Pope, R. and R. Prescott, "Diversification in Relationship to Farm Size and Other Socioeconomic Characteristics." American Journal of Agricultural Economics, Vol. 62, #3 (August 1980): 554-559.

Samuelson, P. "A General Proof That Diversification Pays." Journal of Financial and Quantitative Analysis, Vol. 2 (1967): 1-13.

Sharpe, W. F. "A Simplified Model for Portfolio Analysis." Management Science, (January 1963): 277-293.

Tintner, G. The Variate Difference Method. Bloomington, Ind.: Principia Press, Inc., 1940 (Cowles Commission for Research in Economics, Monograph No. 5), 175 pages.