Florida Journal of Educational Research, 1972, 14

GROUPING FLORIDA SCHOOL DISTRICTS: A FACTOR ANALYTIC STUDY

Martha J. Chang
University of Florida

SUMMARY

Socio-economic, educational, and political data of the 67 Florida counties are examined. First the data are factor analyzed to identify common underlying dimensions. Next, a distance measure was used to set up a matrix of interprofile similarities to be factor analyzed to give types of counties. Thirteen meaningful factors could be substituted for the 74 variables. On transposed analysis, seven type factors were found for the Florida counties.

INTRODUCTION

In classifying school districts within a state or determining school districts for representative sampling, one is often forced to resort to ranking the districts on the basis of one variable. For instance, school districts within an area can be ranked according to per pupil expenditures and then grouped high, medium, and low on this variable. This, however, in no way equates the school districts in a given group on any of the many other variables which might be relevant. Factor analysis enables one to take a multivariate approach to grouping, for it is capable of looking at a number of variables and grouping school districts, counties, urban areas, or other political units on the basis of these variables.

The purpose of this paper is to factor analyze socioeconomic, educational, and political data of the 67 Florida school districts (counties), first grouping the variables to determine common underlying dimensions, and then grouping the school districts. In the terminology of the factor analyst, this study thus focuses on both an R-type analysis and its transpose, a Q-type analysis.

DATA

The 74 variables included in this study were ones which would seem capable of differentiating among the counties in the following socioeconomic, political, and educational areas:

Socioeconomic characteristics
Population characteristics
Employment characteristics
Political characteristics
Population growth and migration
Economic growth
Educational characteristics of the population
School characteristics
Educational support

The data came from the following sources: the United States County and City Data Book, the Florida State Department of Education Research Report No. 73, the Florida Statistical Abstract, and the 1968 election report of the State of Florida. An attempt was made to secure data for the year 1968, although some data necessarily came from the 1960 census. Most of the educational measurements were for the 1967-68 school year. Since there was no desire to produce a factor that reflected mere area size or population, all variables were used as percentages or per capita.

ANALYSIS: COMMON DIMENSIONS OF VARIABLES

As a first step in the analysis of the data, an R-type factor analysis was made, correlating and grouping the 74 variables. (The computer programs used for this study were those developed for the Educational Evaluation Library, University of Florida. These programs are described in Guertin and Bailey, 1970, pp. 293-314). The principal axes method was used. Two iterations brought about satisfactory convergence between the final communality estimates and the row sums of squared loadings. From the factors extracted, 18 were rotated according to the Varimax criterion (Kaiser, 1958). Thirteen of these factors could be interpreted meaningfully; these factors accounted for 78 per cent of the total score variance, 88 per cent of the common variance. Table 1 gives the Varimax rotated factor loadings.

TABLE 1
Varimax Rotated Factor Loadings

Variable	Urban 1	Expenditures Per Pupil 2	Non-White 3	Agriculture 4	Retired 5	Local Support for Schools 6
% fam. with annual income $10,000 or more	.89					
% pop. 25 yrs. & older who completed high sch. or more	.86					
Median sch. yrs. compl. by pop. 25 yrs. & older	.85					
% pop. urban	.85					
% native pop. born in Florida	-.81				-.36	
Personal income per pupil	.81	.32				
Expend. per pupil for auxiliary services	-.80	(.28)				
Per cap. value selected serv.	.79			(-.26)		
% voting for George Wallace	-.78				-.40	
Per cap. personal income	.77					
% pop. rural non-farm	-.78					
% fam. with annual income less than $3000	-.77					
Welfare recipients per 1000 pop.	-.76					
% employed persons in white-collar jobs	.75			-.39		
% pop. increase 1950-60	.74					
% pop. 25 yrs. & older with less than 5 yrs. sch.	-.74		.43			(-.28)
% voting Richard Nixon	.74				.45	
Aver. annual salary paid instructional staff	.69	.56				

TABLE 1
(Continued)

Variable	Urban 1	Expenditures Per Pupil 2	Non-White 3	Agriculture 4	Retired 5	Local Support for Schools 6
% pop. foreign born	.64	.42			.35	
Ratio no. housing units built 1950-60 to no. built before 1950	.64					
% major high sch. classes taught by teachers in field	.64					
% sch. enrol. increase 1957-58 to 1967-68	.63					
% pop. moved into county between 1955 & 1960	.59		(-.26)			
% pop. rural farm	-.57			.46		
Pop. density	.57					
% increase of pop. 65 yrs. & over 1950-60	.56	.30			.43	
% 1968 high sch. grad. entered college in 1968-69	.56					(-.26)
Local sch. revenue per pupil	.56	.46				.53
Pupils per teacher, 7-12	.53	-.32				
Minimum Foundations per pupil	-.49		(.29)			-.34
% instructional personnel below Rank III	-.42	(-.28)	(-.25)			
Per cap. value retail sales	.40					.40
% employed persons in retail sales	.40					.33
% change in sch. membership 1st month 1968-69 1967-68	.35					(.28)
Transfers from out-of-state as % of total enrollment	.37		-.36			
School withdrawals as % of enrol.	.45		-.45			

TABLE 1
(Continued)

Variable	Urban 1	Expenditures Per Pupil 2	Non-White 3	Agriculture 4	Retired 5	Local Support for Schools 6
Per cap. tourist trade, by destination of incoming auto tourists	.31					
Expenditures per pupil for instruction		.89				
Current expense per pupil		.87				
% increase in annual salary paid instruc. staff 1957-58 to 1967-68	.52	.60				
% increase in current expend. per pupil	.31	.54			.33	
Non-exempt assessed val. per pupil		.51			.31	
% instructional staff Rank II or above	.34	.46	(.28)		(.27)	
Millage necessary to meet required local effort for schools		-.43				
Teaching positions as % of academic units earned		.42				
Pupils per teacher elem.	.35	-.31				
Birth rate		(-.28)				
% population nonwhite	(-.27)		.90			
% registered voters Negro			.87			
% voting for Hubert Humphrey	.53		.61			
% males in labor force	-.39		-.44	(.28)		
% teachers men			-.39	(.26)		
% pop. 14-17 enrolled in sch.		-.31				

TABLE 1
(Continued)

Variable	Urban 1	Expenditures Per Pupil 2	Non-White 3	Agriculture 4	Retired 5	Local Support for Schools 6
Per cap. value farm products				.87		
% personal income from agriculture	(-.29)			.86		
% employed persons in agri.	-.47		(.26)	.69		
% elem. teachers taught in field			(.26)	-.58		(.28)
% pop. 65 yrs. & older					.81	
Per cap. mobile homes	(.27)				.77	
% pop. under 18		-.39	(.28)		-.58	
% pop. voting in 1968 pres. elec.		(.29)	-.41		.47	
Local sch. revenue as % of personal income						.78
Local sch. revenue per cap.	.50	.35				.66
% by which taxes exceeded local effort	.36	.39				.59
Per cap. hotel & motel units	(.26)	(.26)				
High sch. enroll. as % of elem.		(.28)				
% pop. 21-24 enrol. in school	(.26)					

TABLE 1
(Continued)

Variable	Enrolled in High School 7	Population Growth 8	Manufacturing 9	University Centered 10	Little Change in Per Cap Income 11	Increase in School Revenue 12	Tourism 13
High sch. enrol. as % of elem.	.77						
Transfers from out of state as a % of total enroll.	-.65						
% pop. 14-17 enrol. in school	.55				.53		
Sch. withdrawals as % of enroll.	-.54						
% pop. rural farm	.36						
Non-exempt assessed val. per pupil	-.30						
Millage necessary to meet required local effort for schools	.34						
Change in Index of Tax-paying Ability from 1958-59 to 1968-69		.85					
% pop. increase 1950-60		.45					
Ratio no. housing units built 1950-60 to no. built before 1950		.44					
% sch. enroll. increase, 1957-58 to 1967-68		.41					
% pop. moved into county between 1957 & 1960		.35					
% voting for Hubert Humphrey		-.31					
% personal income from manufac.			.85				
% employed persons in manufac.			.81				
% instruc. staff below Rank III			-.31				

TABLE 1
(Continued)

Variable	Enrolled in High School 7	Population Growth 8	Manufacturing 9	University Centered 10	Little Change in Per Cap Income 11	Increase in School Revenue 12	Tourism 13
% employed persons in educ. services				.91			
% pop. 21-24 enrol. in school				.80			
% instruc. staff Rank II or above				.41			
% employed persons in white-collar jobs				.33			
% change in per cap. income from 1960 to 1967					-.76		
% major high sch. classes taught by teachers in field					.49		
Per cap personal income					-.36		
% increase in local sch. revenue from 1957-58 to 1967-68						.76	
% pop. moved into county between 1955 & 1960						.34	
% sch. enrol. increase 1957-58 to 1967-68						.34	
Per cap value retail sales		(.27)				-.32	
Per cap. hotel & motel units							.71
Per cap. tourist trade, by destination of incoming auto tourists							.76
Welfare recipients per 1000 pop.	(.26)						
Population density	(.28)						
% increase of pop. 65 & over 1950-60		(.27)					
% change in sch. membership 1st month 1968-69 over 1st month 1967-68	(-.27)						
% males in labor force			(.29)				

While this study was in its early stages, a report of a related study by Robert S. Stephenson and Jacob G. Beard appeared (1971). Using different data, the latter focused entirely upon the dimensions of the variables, not upon the grouping of the counties. Of the 46 variables in the Stephenson-Beard study, 23 were common to this present study. These common variables represent 61 per cent of the variables in the Stephenson-Beard study and 39 per cent of the ones in the present study. Six of the seven factors in the former study are similar to those in the present study, as can be seen by comparing the lists of factors in Table 2. Of the factors in the Stephenson-Beard study, only Community Size cannot be found in the present study, and this, of course, because size variables were not included in the study. The use of more broadly based variables resulted in seven additional factors in the present study and in clarification of some of the six factors common to both studies.

TABLE 2
Comparison of Factors

Stephenson-Beard Study		Present Study	
Factor	Percentage of Total Score Variance	Factor	Percentage of Total Score Variance
1 Community Financial & Educational	29.0	1 Urban (High Socioeconomic)	25.9
2 Community Growth	11.2	2 Expenditures Per Pupil	7.5
3 Expenditure Per Pupil	9.4	3 Nonwhite	5.4
4 School Holding Power	6.3	4 Agriculture	5.3
5 Community Size	5.6	5 Retired	5.3
6 Local School Support	4.3	6 Local Support for Schools	4.2
7 Minority Group	4.3	7 Enrolled in High School	4.2
		8 Population Growth	3.4
		9 Manufacturing	3.4
		10 University Centered	3.3
		11 Little Change in Per Capita Income	3.0
		12 Increase in School Revenue	2.5
		13 Tourism	2.5

An example of such clarification may be found in comparing Factor 7 of the present study, Enrolled in High School, with the Stephenson-Beard Factor 4, School Holding Power. The latter label implies that the school environment in the school districts which contribute variance to this factor are such that the drop-out rate is lower. The broader socioeconomic variables used in this present study gives a different picture, pointing to out-of-school environment as a greater influence than any power of the schools to hold students in school. For instance, there seems to be a relationship between high school enrollment and a shift in the population to an older age bracket. Although the variables "High school enrollment as a percentage of elementary," and "Percentage enrollment as a percentage of elementary," and "Percentage of the population 14 to 17 enrolled in school," seem on the surface jointly to be the key to this factor, there is an intercorrelation of only .31 between the two. In other words, less than ten per cent ($.31^2$) of the variance of one can be explained by the other. Slightly larger variance of "Percentage of the population 14 to 17 enrolled in school" is associated with negative correlations with such socioeconomic variables as "Non-exempt valuation per pupil" (-.51). "Percentage change in per capita income" (-.39), "Percentage of employed persons in agriculture"(-.38), and "Per capita income" (-.35), indicating that some of the variance in this variable seems to be related to lack of economic opportunities for youth. The low correlations with school variables do not point to a tie between a lowered drop-out rate and the holding power of the school.

Beyond this type of clarification, the present study, using data for a different school year, provides a validation for the earlier study. The seven additional factors are the result of wider ranging variables, valuable for grouping the school districts in areas not included in the Stephenson-Beard study.

ANALYSIS: TYPES

The conventional method for transposing a matrix of scores is to standardize the scores for each case and then intercorrelate the scores for each case with those for each other case. The principal axes method is then used, as in R-type analysis, to determine dimensions or factors within the intercorrelation matrix. This method groups cases on the basis of the shape of their profiles while ignoring level or magnitude of scores. In analysis of the characteristics of political units, magnitude of scores is more important than the shape of profiles. Therefore, it is not surprising

that an attempt to use the conventional correlational index for these data brought about no meaningful grouping of counties. A grouping that proved much more meaningful was one based on a distance measure as developed by Cronbach and Gleser (1953). The index, known as d, is simply the square root of the sum of the squared distances between two profiles across all variables. The d's are then transformed into a matrix of indices of similarity, with values approaching 1.0 indicating greatest similarity and those closest to 0.0 least similarity. (For an explanation of this transformation, see Guertin and Bailey, 1970, p. 269.) This matrix is then used the same way that an intercorrelation matrix would be used, with the highest coefficients employed as the original communality estimates.

The resulting 35 principal axes account for 91 per cent of the total score variance. Seven factors, accounting for 81 per cent of the total score variance, 90 per cent of the common variance, were rotated both orthogonally to the Varimax criterion and obliquely using the Simple Loadings procedure (Jennrich and Sampson, 1966). With over half of the intercorrelations between the simple loading primary factors above .35, the oblique solution was preferred. This solution is shown in Table 3.

In order to determine the nature of each group of counties, extensive hand graphing was done comparing counties within the groups and contrasting counties from different groups. Some graphing was done using all of the 74 variables, but for most part graphing was done on the basis of factor score estimates of the counties on the 13 interpretable R factors. The reason is obvious--graphing 13 points as opposed to the tedious graphing of 74 points.

Counties with high loadings on Group A tend to have high scores on the socioeconomic variables, and the counties which load highest on this group have high scores also on such factors as Retired, Population Growth, Community Support for Schools, and Tourism. These are the tourist counties, the boom counties. Moving further down this group to counties with lower but still significant loadings, we find counties which have scores on these factors which are closer to the mean, or counties which have relatively high scores on some of these factors mentioned but not on others, e.g., counties which do not have the urban, high socioeconomic characteristics, but which nevertheless have an economy significantly tied to tourism.

Group B represents counties at the other extreme, that is, counties which have very low scores on socioeconomic characteristics. These are the northern counties which have few tourists, few retirees, and relatively smaller population and economic growth. Group B may be contrasted also with Group F. Both are in the northern tier and both represent socioeconomic areas which are below the mean, but the counties in Group B have a lower socioeconomic level while those in Group F have a larger percentage of non-white population. It is interesting that the state's poorest counties represent a distinct grouping from those counties which have the highest population of nonwhite.

Counties in Group C are those which depend upon agriculture to a larger degree than do the other counties. By contrast, Group D includes counties which depend more upon manufacturing and much less upon agriculture. Alachua and Leon, the two counties loading on Group E, and the home of the University of Florida and Florida State University, respectively, have very high scores on the factor University Centered, and have scores on all other factors which are very close to each other. No interpretation for Factor G was satisfactory.

Table 3 portrays a grouping of the counties on the basis of the entire 74 variables. There may be additional insight gained from looking at groupings on the basis of the four specific educational factors: Expenditures Per Pupil, Local Support for Schools, Enrolled in High School, and Increase in School Revenue. Grouping on the basis of factor score estimates, though still only a ranking procedure, is certainly preferable to using scores on just one variable. For instance, the factor Expenditures Per Pupil is more complex than the variable "Expenditures per pupil" because the former takes into account correlations between a number of variables, including demographic data which may help to explain some expenditure variables.

Rather than show a ranking of the entire 67 counties on factor scores on these four factors, Table 4 includes only those counties which have factor scores larger than 1.0. Thus only the extremes are shown. Any interpretation of these rankings should be done with care. For instance, the reasons why one county spends more per pupil than another county may be related to a number of things--ratio of school age children to adults, population density (a sparsely populated area frequently results in smaller classes per teacher and higher transportation costs), average salary in the community, and other relevant factors.

TABLE 3
Oblique Rotated Factor Loadings

County	A	B	C	D	E	F	G
Sarasota	.94						
Palm Beach	.84						
Pinellas	.80						
Charlotte	.79						
Volusia	.72						
Broward	.70						
Dade	.66						
Manatee	.60						
Lake	.59		.31				
Indian River	.57						
Collier	.53						.31
Lee	.53		.32				
Orange	.50		.31				
Hillsborough	.48			.35			
Highlands	.48		.35				
Martin	.47		.33				.31
St. Lucie	.44		.35				
Osceola	.42		.39				
St. Johns	.40						
Hernando	.40						.35
Duval	.38			.34			
Monroe	.36		.34				
Marion	(.29)					(.29)	
Holmes		.99					
Washington		.80					
Walton		.76					
Lafayette		.73					
Calhoun		.69					
Gilchrist		.68					
Jackson		.59				.34	
Liberty		.56					
Suwannee		.52					
Baker		.50	.30				
Wakulla		.46					.33
Sumter		.46					
Union		.40					
Levy		.39					
Bradford		.38					
Franklin		.33					
Hardee			.76				
Okeechobee			.68				
Hendry			.58				
Clay			.55				
Okaloosa			.55		.37		
De Soto			.53				
Pasco	.38	.31	.47				
Seminole			.44				
Polk	.34		.38				
Taylor				.70			
Gulf				.69			
Nassau				.55			
Putnam				.52			
Dixie		.33		.52			.31
Santa Rosa				.47			
Escambia				.38			
Bay				.35			
Brevard				.33			
Leon					.99		
Alachua					.95		
Gadsden						.70	
Jefferson						.66	
Madison		.41				.54	
Columbia		.31				.37	
Flagler						.30	
Glades							.47
Citrus	.45						.45
Hamilton						.32	.39

CONCLUSIONS

An R-type factor analysis of 74 socioeconomic, political and educational variables of the 67 Florida county-school districts yielded 13 interpretable orthogonal factors, six common to an earlier study. The seven additional factors resulted from the more broadly based variables used, but these same variables helped to explain ramifications of some of the six factors common to both studies as well.

The focus of this study, however, was on a Q-type analysis, grouping the counties. For this purpose a distance analysis proved more useful than the conventional correlational analysis. Seven groupings resulted. By hand graphing the counties in each group on the basis of factor score estimates on the R factors, group characteristics were determined and described.

These groupings have potential usefulness in determining school districts for representative sampling, providing matching school districts for experimental and control designs, and for similar research needs. Probably more fundamental, such groupings, along with the R-type analysis, provides a promising empirical means of making sense out of the complex multidimensional world of which the schools are a part.

REFERENCES

Bureau of Economic and Business Research, College of Business Administration, University of Florida, Florida statistical abstract, 1969. Gainesville: 1969.

Cronbach, L. J., and Gleser, C. G. Assessing similarity between profiles. Psychological Bulletin, 1953, 50 456-73.

Guertin, Wilson H., and Bailey, J. P., Jr. Introduction to modern factor analysis. Ann Arbor, Michigan: Edwards Brothers, 1970.

Jennrich, R. I., and Sampson, P. F. Rotation for simple loadings. Psychometrika, 1966.

Kaiser, H. F. The varimax criterion for analytic rotation in factor analysis. Psychometrika, 1958, 23, 187-200.

State of Florida, Department of Education. Research Report No. 73, Ranking of the counties . . . 1968. Tallahassee, 1968.

State of Florida, Secretary of State. Tabulation of official votes cast in the general election, November 5, 1968.

Stephenson, Robert S., and Beard, Jacob G. Common dimensions of the school, social, and economic environment in Florida: An empirical study. Florida Journal of Educational Research, 1971, 13, 49-57.

U. S. Bureau of the Census, County and city data book. Washington, D. C.: U. S. Government Printing Office, 1962.