Similarities Among the State Highway Systems

DAVID T. HARTGEN

Results are reported for a recent study to identify similarities among the financial, infrastructure, and operational structures of the state highway systems, and to group states according to these similarities. The purpose of the study was to classify states according to similar systems and problems, rather than according to closeness or geography. An additional goal was to determine the degree to which states that are widely separated may have similar problems, thereby suggesting political alliances for various issues. A structure for classifying state highway agency financing is first proposed and described. Then, a data base containing 61 data items for each state is developed using cross-sectional data from 1984, 1985, and 1986. The source of these data items is primarily *Highway Statistics*, supported by organization studies in the transportation literature. Data consist of measures of size, road and bridge condition, taxes, revenues, disbursements, and agency characteristics. After a review of descriptive rankings of states on a number of key variables, the data are then factor-analyzed using varimax rotation. This procedure yields a smaller number of data items found to most clearly separate these states. States are then clustered according to their ratings on these variables. Clusters based on the full data set containing 61 variables are compared with clusters based on reduced data sets containing 19, 13, and 7 variables. Results show that with 19 variables, the state highway systems can be quite accurately clustered into five distinct groups as follows: (1) Alaska; (2) the far western states, along with Michigan, Hawaii, Florida, West Virginia, and northern New England; (3) southern New England and seaboard Middle-Atlantic; (4) midwestern and southern states, including Washington; and (5) very large states (California, Texas, New York, and Pennsylvania). Within each of these groups, a number of subgroupings identify strong regional and content coalitions. The group structure appears to be reasonably robust under a variety of assumptions, with regional subgroupings particularly strong for New England, the southeastern seaboard, the far west, and big states. It is concluded that while nearest neighbor geographies may be useful for many political liaisons, state highway agencies should also look at broader similarities with states that are not immediate or nearest neighbors.

It is well understood that the problems associated with planning, financing, building, and maintaining state highway systems are extensive and complex. Each of the state highway agencies is faced with many problems involving each of these activities. These problems are generally similar from state to state, but their details vary immensely because each state is unique in its location, context, and capabilities. Thus, under the general guidance of federal law, state policy, and goal structure, each of the states operates more or less independently in managing its state highway system.

In spite of these differences, there also are many similarities in the highway management process: each state operates within the general guidance of federal law; each state has organized and maintains a system of state-owned highways and an agency responsible for them; each state has developed and relies upon its own funding sources for revenues necessary to maintain its system. In theory, therefore, while many differences in procedures and problems separate the states, we should also recognize that many similarities bind them as they deal with these complex issues.

A fundamental hypothesis worthy of testing is whether the state highway agencies have problems and solutions similar to their closest neighbors, or whether the similarities extend beyond the borders of immediate states to those further away. One might expect that geographical, historical, political, climatic, and development conditions would lead to the greatest similarities being shared by the states closest to each other. Conversely, as the Interstate System has fostered a larger proportion of regional traffic, and as state highway agencies share technology and procedures through technical as well as political processes, one might also expect that these geographically based similarities are declining over time in favor of similarities of problem and solution. If this is so, then state liaisons based solely upon geography are likely to be declining relative to those based upon system similarity. In a nutshell, the purpose of this paper is to review the extent to which geographical or structural similarities among the state highway agencies can be utilized to identify such coalitions. Our purpose is to determine whether the nature of these coalitions is primarily geographic or structural.

THEORY AND METHOD

While each of the state highway agencies is responsible for planning, constructing, and maintaining its highway system, in addition to other duties, each state is largely independent in developing the organizational and fiscal structure necessary to achieve those goals. In addition, the states may be expected to vary widely along the key dimensions relating to these activities. These key dimensions are as follows:

- geography, including soil conditions, climate, freeze-thaw cycles, and weather;
- size and extent, including size of the highway system for which the agency is responsible, measures of area and population size for the state itself, and measures of traffic;
- system condition, specifically congestion, bridge and highway condition, and performance measures;
- sources of revenue, particularly tax rates for different fuels, vehicles, licenses, and other sources;

Department of Geography and Earth Sciences, University of North Carolina at Charlotte, Charlotte, N.C. 28223.

- agency characteristics, particularly measures of agency size, focus, and structure;
- expenditures, particularly the magnitude and density of expenditures per mile; and
- network access and measures of network accessibility relative to the area and population size.

Figure 1 shows a simple model that describes how most states deal with the problem of financing highway systems (*1–3*). Within the context of a general goal direction and a geographic environment, the state sets tax rates to generate revenue needed to deal with the highway problems or conditions. Expenditures then result in highway improvements that, when compared with goals, allow for revisions of revenue streams. Within legislative, managerial, and operational environments, the process goes along more or less continuously. That is, periodic readings of highway conditions and needs are used to identify requirements, which in turn lead to revenue generation through taxing actions. Over time, system improvements result in changes in condition, which modify needs requirements.

The structure described in Figure 1 can be represented as a series of simultaneous equations in which data for a number of years is used as the basis for model development. For instance, let:

g = geographic attributes,
G_y = goals vector, year y,
T_y = tax rates,
S_y = size measures,
Rev_y = revenue,
$Needs_y$ = needs,
$Cond_y$ = conditions,
$Disb_y$ = disbursements,
$Maint_y$ = maintenance expenditures, and
Cap_y = capital expenditures.

Then, a system of simultaneous equations can be developed to describe this process. For instance, the relationship in Figure 1 can be expressed as:

$$T_y = f(G_{y-1}, N_{y-1}),$$
$$Rev_y = f(T_{y-1}, Size_y, Disb_{y-1}),$$
$$Cond_y = f(Size_y, Cap_{y-1}, Maint_{y-1}),$$
$$Needs_y = f(Cond_{y-1}),$$
$$Disb_y = f(Rev_{y-1}, Cond_{y-1}),$$
$$Cap_y = f(Disb_y), \text{ and}$$
$$Goals_y = f(Cap_y, Maint_y, Needs_y).$$

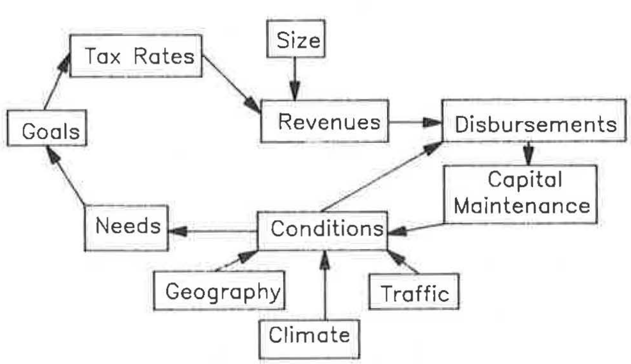

FIGURE 1 Simplified state highway financing model.

The specific functional form of these models would be determined by two-stage least squares calibration. Generally, linear models are used.

Models of this structure are particularly difficult to develop and calibrate because they require data on systems expenditures and performance for a number of time periods. Recognizing the limitations of a simpler design, we have chosen to use primarily cross-sectional data for model estimation purposes.

Numerous studies contain comparative cross-sectional (and occasionally time series) data on highway financing. The most extensive and comprehensive is *Highway Statistics*, published since the 1920s. This series contains aggregate data, by state, on many aspects of financing, road conditions, and traffic. Its primary shortcomings are in agency employment, which it does not contain, and in "quality" measures of system performance (such measures as congestion and pavement condition were only recently added) (*4*). These reports can be analyzed over time, but the Federal Highway Administration (FHWA) did not (until recently) keep its data in that fashion, so comparisons are difficult. The summary document, *Highway Statistics: Summary to 1985*, contains five- and ten-year trend statistics for key indicators, by state; and a few data items are kept annually (*5*). An earlier document contains state-level trend data for the years 1957–1975, by modal programs (*6*). The series *Highway Taxes and Fees* (1981, 1984, 1987) describes procedures needed by the states to collect and disburse taxes; a companion document, *Financing Federal Aid Highways*, describes the federal process (*7,8*). Other documents contain data on bridge statistics and agency size (*9,10*).

Reports prepared for fiscal reviews by individual state highway departments also typically contain general comparisons with other states deemed "similar," but generally no detailed comparative analysis of states are undertaken. Apparently, comparative reviews of state highway funding practices are not as common a subject of investigation as they once were. Of 24 studies reviewed for this paper, most were conducted before 1970. An NCHRP Project (1970) reviewed state level budgeting practices, relying on states' reports from the field. In one of the very few comparative analytical studies of state highway financing, Phelps (*1*) developed a stock-adjustment model based on economic theory to describe the timing of state and local highway capital outlays; data were time series for 1951–66, primarily from *Highway Statistics*. An even earlier study (*3*) compared the collection and distribution practices of motor vehicle revenues in 34 states, and concluded that the primary uses (71 percent) were for state highway construction and maintenance. This detailed paper, written in 1927, contains 1925 data, perhaps commenting indirectly on the speed of today's "modern" data delivery systems. Rao (*2*) compared the funding practices and procedures of the states, focusing on bonding and other revenue-building approaches, but did not analyze the similarities of states statistically. None of the documents reviewed used such tools as factor analysis or cluster analysis. The author is not aware of any such applications to highway data, although at least one study (*11*) attempted a path analysis of transit property performance statistics. In general, the subject is unresearched.

One way to determine the similarities among state financing and organization structure is to calibrate a model such as the above, using time series data for each state. A comparison of the coefficients of the models for each state would then lead to rejection or acceptance of the null hypothesis that the states are homogeneous in their financing structures. However, such a methodology would be extremely consuming in both time and effort. Therefore, a short approach relying primarily on cross-sectional data is used instead. This approach relies heavily on the cross-sectional data available for each of the state highway agencies, primarily reported in *Highway Statistics*, supplemented by data on transportation agency employment (*9*) and data on bridge conditions (*10*).

Data representing primarily 1984, but also some 1980, 1983, 1985, and 1986 data, were consolidated from these sources into a state level data base using spreadsheet (EXCEL) and SAS environments. The specific variables used in the analysis consist of 61 data items drawn from the above sources and grouped into several major classifications. These data items, extracted in Table 1, contain information on the following dimensions: (1) system size and state size measures, (2) measures of condition, (3) fiscal sources and revenues, (4) expenditures, (5) agency size, (6) expenditure efficiency.

TABLE 1 MEASURES OF STATE HIGHWAY AGENCY STRUCTURE

Category	Subclass	Variable	Source	Name
Size	Mileage	• Rural miles under State control	(1)	RUSC
		• Urban miles under State control	(1)	UUSC
		• Total miles under State control	(1)	TOTUSC
		• Total state mileage	(1)	TRUM
		• Rural Interstate mileage	(4)	RITOT
		• Rural principal arterial mileage	(4)	RPATOT
		• Urban Interstate miles	(4)	UVCTOT
		• Urban principal arterial mileage	(4)	UPATOT
	Traffic	• Annual VMT (million)	(2)	VMT
	Population	• Resident population, 1983	(8)	RESPOP
	Drivers	• Licensed drivers/1000 pop.,1983	(8)	LDRP
	Vehicles	• Registered motor vehicles/ 1000 pop., 1983	(8)	RVMPOP
		• Licensed drivers/reg. motor veh.	(8)	LDRMV
	Bridges	• Federal aid number of bridges	(9)	FANOBR
		• Non-Fed. aid number of bridges	(9)	NFANOBR
		• Total number of bridges	(9)	NOBRIDGE
Condition	Congestion	• Miles of rural Interstate with V/C.7-.95	(4)	RVC1
		• Miles of rural Interstate with V/C >.95	(4)	RVC2
		• Miles of rural principal arterial with V/C.7-.95	(4)	RPAVC1
		• Miles of rural principal arterial with VC>95	(4)	RPAVC2
		• Miles of Urban Interstate with V/C.7-.95	(4)	UVC1
		• Miles of Urban Interstate with V/C>.95	(4)	UVC2
		• Miles of Urban Principal Arterial with V/C .77-95	(4)	UPAVC1
		• Miles of Urban Principal Arterial with V/C>.95	(4)	UPAVC2
		• Interstate + Arterial Networks % of roads with V/C>.77		PVC
	Pavement	• Rural Interstate miles PSR<2.5	(5)	RIPSRMI
		• Urban Interstate miles PSR<2.5	(5)	RIPSRMI
		• Interstate % PSR<2.5	(5)	PSR
	Bridges	• Number of FA deficient bridges	(9)	FADEF
		• Number of non-FA deficient bridges	(9)	NFADEF
		• Percent of deficient bridges (FA & NFA)	(9)	DEFBR
		• Percent of FA deficient bridges		PDFABR
		• Percent of non-FA deficient bridges		PDNFABR
Taxes	Gasoline	• Gasoline tax rate cents/gal	(3)	GASTAX
	Diesel	• Diesel tax rate	(3)	DIESLTAX
	Others	• Special fuels rate	(3)	OTHERTAX
Revenues	Fuels	• Receipts for State hwy.-motor fuels	(6)	FUELTAX
	Vehicles	• Receipts for State hwy.-vehicle/carriers	(6)	MVMCTAX
	Total	• Total receipts		TOTREV

TABLE 1 (*continued on next page*)

TABLE 1 (*continued*)

Category	Subclass	Variable	Source	Name
Disbursements	Capital	• Capital outlay for roads & bridges	(7)	RBCORB
	Maintenance	• Outlay for maintenance & traffic services	(7)	SERVMTS
	Administrative	• Outlay for administrative & misc.	(7)	MISCADM
	Total	• Total disbursements	(7)	TOTDISB
Employees	Number	• No. of engineers in management, 1984	(10)	MGMTENG
		• Total no. of engineers, 1984	(10)	TOTENG
		• Total no. of planners	(10)	PLANRS
		• Total no. of computer employees	(10)	CMPTR
		• Total employees, 1980	(10)	TOTEMP80
		• Total employees, 1984	(10)	TOTEMP84
		• % change in employees 1980-84		CHGEMP
	Salaries	• District Engineer's salary, 1984, $	(10)	DE
		• Project Engineer's salary, 1984, $		PE
		• CE graduate's salary, 1984, $		BSCE
Rates				
	Staff density	• Miles of road under State control per employee		MIPEREMP
	Expenditure density	• Capital outlay for roads and bridges per mile		CORBPUSC
		• Maintenance & traffic services expenses per mile		MTSPUSC
		• Administrative expenses per mile		ADMPUSC
		• Total disbursements per mile		DISBPUSC
		• Total disbursements per population		DISPOP
		• Total disbursements per motor vehicle		DISBPRMV
	Expenditure efficiency	• Capital outlay for road & bridges per employee		CORPEMP
		• Maintenance & traffic services expenses per employee		MTSPEMP
		• Administrative expenses per employee		ADMPEMP
	Engineering focus	• Engineers per employee		ENGPEMP
		• Mgt. engineers per employee		MGTEPEMP

Notes

(1) Highway Statistics, 1984, Table HM10
(2) Highway Statistics, 1984, Table VM2
(3) Highway Statistics, 1984, Table MF1
(4) Highway Statistics, 1984, Table HM61
(5) Highway Statistics, 1984, Table HM62
(6) Highway Statistics, 1984, Table SF-3
(7) Highway Statistics, 1984, Table SF-4
(8) Highway Statistics, 1984, Table SF-17
(9) Sixth Annual Bridge Report, 1984 data, Table 10
(10) TRB Special Report No. 207, 1984 data, Table 2-16.

The methodology for analyzing this data is relatively straightforward. It consists of three basic steps:

1. Descriptive statistics were prepared for key measures of state systems. These ranked data streams were used to develop a preliminary overview of relative state groupings.
2. Using factor analysis techniques and varimax rotation, the descriptive statistics were factor analyzed to develop a reduced set of variables representative of each factor group.
3. Using these variables, the states were clustered along similarity lines using Ward "nearest neighbor" clustering algorithms. Reduced variable structures were identified for 7, 13, 19, and a total of 61 variables. The SAS procedure CLUSTER was used for this approach. Separate cluster analyses were then prepared for large and small states, to determine whether state size is critical in the grouping process. Characteristics of each cluster were then prepared.

RESULTS

Size Measures

Table 2 shows various measures of size for these states. California, New York, and Texas lead the list in population, vehicles, drivers, and VMT, with Pennsylvania, Illinois, Ohio, and Florida close behind. With respect to highway system extent, Georgia, North Carolina, Texas, and Virginia have the largest state highway systems, all over 50,000 miles.

Measures of Condition

Network condition can be measured according to a number of criteria. Data readily available include the percent of mileage with volume-to-capacity ratios greater than .77, the per-

TABLE 2 MEASURES OF SIZE

OBS	STATE	POP	VEHICLES	DRIVERS	VMT	MILES	BRIDGES
1	CA	25044.6	17781.7	16714.8	196537	18213	22260
2	NY	17713.1	8421.0	9599.9	87268	16394	17419
3	TX	15758.2	11734.8	11500.1	137737	71448	44036
4	PA	11896.0	6843.6	7459.6	74297	44000	21719
5	IL	11484.8	7507.1	6981.6	69910	17609	25058
6	OH	10725.9	7778.3	7389.4	74895	20221	28969
7	FL	10686.9	8812.0	8371.4	85475	11536	10005
8	MI	9109.4	6300.6	6363.7	63470	9510	10401
9	NJ	7473.7	4939.4	5433.3	52312	3167	5732
10	NC	6094.6	4615.0	3968.9	48182	76920	15712
11	MA	5786.7	3835.7	3682.3	38537	3613	4776
12	GA	5713.3	4206.5	3743.8	50486	86655	14191
13	VA	5535.7	3890.0	3695.5	44527	54782	12493
14	IN	5506.3	3837.7	3530.7	41074	11344	17682
15	MO	4949.9	3433.4	3330.4	38535	32317	23726
16	WI	4752.4	3212.9	3084.4	35367	12519	12822
17	TN	4672.2	3548.0	2944.9	36523	11171	18110
18	LA	4418.7	2872.1	2757.3	31588	16419	14215
19	MD	4304.5	3015.2	2804.1	31702	5239	4059
20	WA	4304.3	3339.2	2871.7	34248	18525	6796
21	MN	4150.4	3288.6	2367.8	31826	13443	12906
22	AL	3953.2	3147.0	2391.7	32961	11688	15513
23	KY	3710.9	2621.8	2202.3	27951	25120	12484
24	OK	3304.2	2775.6	2192.7	30981	13056	22021
25	SC	3263.6	2052.9	2011.8	25971	40338	8890
26	CO	3151.2	2641.4	2218.8	24588	9301	7147
27	CT	3131.4	2308.8	2262.6	21076	3896	3724
28	AZ	2955.5	2294.8	2180.1	20613	5786	5140
29	IA	2910.6	2474.9	1930.4	20497	10160	26112
30	OR	2655.2	2117.2	1905.5	20943	10856	6869
31	MS	2587.6	1561.3	1811.1	18442	10324	16728
32	KS	2429.0	2050.2	1681.1	18717	10692	25656
33	AR	2322.2	1444.9	1647.2	16621	16111	14336
34	WV	1967.5	1295.2	1411.8	12671	31356	6608
35	UT	1622.5	1075.6	925.0	11661	5584	2371
36	NE	1594.7	1232.7	1097.1	11968	10385	16197
37	NM	1401.2	1238.3	767.7	12432	12406	3420
38	ME	1148.0	766.9	774.6	9345	7999	2592
39	HI	1026.7	617.0	573.8	6505	1059	1038
40	ID	987.6	877.9	649.6	7768	5085	3622
41	NH	958.7	801.8	697.6	7294	4398	2551
42	RI	952.7	598.7	604.7	5300	1952	689
43	NV	890.1	730.6	679.4	7332	5183	1011
44	MT	816.2	828.3	488.7	7386	7830	4777
45	SD	698.6	629.3	484.5	6401	7896	7061
46	ND	689.2	675.5	439.1	5377	7304	5475
47	DE	605-5	427.0	431.2	5138	4616	691
48	VT	523.6	367.6	360.3	4403	2787	2654
49	WY	514.3	502.4	391.9	5127	6622	2851
50	AK	478.6	350.6	287.5	3589	11426	835
51	DC	.	.	.	3214	1102	236

cent of roads in poor condition, and the percentage of deficient bridges. Table 3 shows these statistics, and indicates that the District of Columbia, Missouri, Connecticut, Kansas, and West Virginia lead these lists. The District is, of course, particularly high on congestion.

Fiscal Sources

A simple measure of state gasoline tax rates, gas tax and diesel tax, shows that Washington has the highest gasoline and diesel tax combination, followed by Minnesota, Louisiana, Montana, and Wisconsin (Table 4).

Expenditures

Table 5 shows that with respect to disbursements, a number of southern states, particularly North Carolina, South Carolina, Georgia, and Virginia, have the highest mileage per dollar expenditures, while several eastern states, including Maryland, Connecticut, Massachusetts, and New Jersey, have the lowest rates.

Overall Status

A useful measure of overall status is the "misery index." This is constructed as a combination of condition (percentage of deficient bridges, deficient pavement, and congested roads), gasoline and diesel taxes, and miles of road per dollar available. High ratings on the "misery index" mean that a state has many problems (per mile) and not much money (per mile) to deal with them. Table 5 also rates states according to the "misery index" and shows that a group of southern states scored particularly high on this index, primarily because of their very high mileage per dollar expended. These states include North Carolina, South Carolina, Georgia, West Vir-

TABLE 3 STATES RANKED BY ROAD CONDITION

OBS	STATE	OVERALL CONDITION	PERCENT V/C	POOR CONDITION	PERCENT DEFICIENT BRIDGES
1	DC	1.41	0.88	0.33	0.19
2	MO	1.01	0.06	0.27	0.68
3	CT	0.91	0.31	0.00	0.60
4	KS	0.86	0.01	0.28	0.56
5	WV	0.84	0.12	0.14	0.58
6	MD	0.82	0.31	0.15	0.36
7	MT	0.82	0.00	0.10	0.70
8	MS	0.80	0.01	0.17	0.61
9	NY	0.79	0.12	0.00	0.66
10	NC	0.76	0.07	0.03	0.65
11	OK	0.75	0.04	0.15	0.54
12	RI	0.73	0.21	0.31	0.20
13	NJ	0.72	0.37	0.01	0.34
14	WI	0.68	0.02	0.16	0.49
15	MI	0.68	0.10	0.23	0.33
16	NH	0.67	0.14	0.00	0.52
17	CO	0.67	0.08	0.16	0.42
18	LA	0.66	0.09	0.04	0.52
19	HI	0.65	0.38	0.00	0.27
20	PA	0.65	0.19	0.10	0.34
21	AR	0.63	0.01	0.01	0.60
22	NE	0.62	0.00	0.00	0.62
23	CA	0.62	0.26	0.08	0.27
24	IN	0.62	0.02	0.00	0.58
25	VT	0.61	0.04	0.04	0.52
26	TN	0.60	0.07	0.01	0.51
27	KY	0.59	0.12	0.05	0.41
28	MA	0.58	0.25	0.00	0.33
29	AL	0.58	0.03	0.00	0.54
30	ND	0.58	0.00	0.00	0.57
31	IA	0.57	0.00	0.02	0.54
32	GA	0.56	0.10	0.00	0.45
33	TX	0.54	0.10	0.04	0.40
34	AK	0.53	0.04	0.29	0.19
35	ID	0.53	0.00	0.18	0.34
36	WA	0.51	0.24	0.00	0.26
37	NM	0.49	0.00	0.28	0.20
38	VA	0.49	0.09	0.05	0.34
39	IL	0.47	0.06	0.07	0.32
40	SD	0.46	0.01	0.00	0.45
41	OH	0.40	0.10	0.08	0.21
42	MN	0.39	0.04	0.03	0.32
43	ME	0.39	0.03	0.04	0.31
44	DE	0.37	0.21	0.00	0.16
45	FL	0.32	0.03	0.00	0.29
46	WY	0.30	0.00	0.03	0.26
47	SC	0.30	0.04	0.01	0.23
48	OR	0.29	0.06	0.00	0.22
49	UT	0.28	0.07	0.02	0.18
50	NV	0.19	0.01	0.01	0.16
51	AZ	0.14	0.02	0.04	0.07

*Condition = Percent v/c > .77 + Percent PSR < 2.0 + Percent Deficient Bridges

ginia, Missouri, and Virginia. At the bottom of the list, states with good conditions, low tax rates, and high expenditures per mile of system, include Nevada, Utah, Wyoming, Arizona, and Florida.

Factor Structure

Tables 6 and 7 show results of the factor analysis of these 61 variables. Results show that approximately six factors are sufficient to describe the data structure. These factors are described as follows: (1) size and congestion, (2) expenditure per mile of system, (3) total mileage and fuel revenues, (4) number of bridges, (5) deficient bridges, and (6) tax rates, pavement conditions, and expenditures. Consequently, the model structure suggested here assumes that size, tax rates, bridge deficiencies, and expenditures sufficiently distinguish between the state fiscal patterns.

Cluster Analysis

To undertake the cluster analysis, variables were selected from each of the factor groups described above in decreasing numbers. The full cluster analysis contained all 61 data items, while reduced cluster analyses contained 19, 12, and ultimately 7 data items, respectively. The specific items selected for each cluster analysis were chosen based on factor loadings, internal independence, and representativeness for each factor group. Table 6 shows the specific variables used for each cluster run. States were grouped into large and small classifications on the basis of the total number of miles under state

TABLE 4 STATES RANKED BY TOTAL FUEL TAX RATE (1984)

OBS	STATE	TAXES	GAS TAX	DIESEL TAX	OTHER TAX
1	WA	36.0	18.0	18.0	.
2	MN	34.0	17.0	17.0	.
3	LA	32.0	16.0	16.0	.
4	MT	32.0	15.0	17.0	.
5	WI	32.0	16.0	16.0	.
6	DC	31.0	15.5	15.5	.
7	CT	30.0	15.0	15.0	.
8	MI	30.0	15.0	15.0	.
9	NE	29.4	14.7	14.7	.
10	ID	29.0	14.5	14.5	.
11	IA	28.5	13.0	15.5	13
12	ME	28.0	14.0	14.0	.
13	NH	28.0	14.0	14.0	.
14	UT	28.0	14.0	14.0	.
15	AL	27.0	13.0	14.0	.
16	MD	27.0	13.5	13.5	.
17	VT	27.0	13.0	14.0	.
18	IL	26.5	12.0	14.5	12
19	AZ	26.0	13.0	13.0	.
20	ND	26.0	13.0	13.0	.
21	RI	26.0	13.0	13.0	.
22	SC	26.0	13.0	13.0	.
23	SD	26.0	13.0	13.0	11.0
24	CO	25.0	12.0	13.0	.
25	NC	24.0	12.0	12.0	.
26	NV	24.0	12.0	12.0	.
27	OH	24.0	12.0	12.0	.
28	PA	24.0	12.0	12.0	.
29	TN	23.0	10.0	13.0	9.0
30	IN	22.2	11.1	11.1	.
31	DE	22.0	11.0	11.0	.
32	KS	22.0	11.0	11.0	10.0
33	MA	22.0	11.0	11.0	6.9
34	NM	22.0	11.0	11.0	.
35	VA	22.0	11.0	11.0	.
36	WV	21.0	10.5	10.5	.
37	AR	20.0	9.5	10.5	7.5
38	KY	20.0	10.0	10.0	.
39	TX	20.0	10.0	10.0	.
40	FL	19.4	9.7	9.7	.
41	MS	19.0	9.0	10.0	8.0
42	CA	18.0	9.0	9.0	6.0
43	NY	18	8.0	10.0	8
44	OK	18	9.0	9.0	.
45	OR	18	9.0	9.0	.
46	HI	17	8.5	8.5	6
47	AK	16	8.0	8.0	.
48	NJ	16	8.0	8.0	4
49	GA	15	7.5	7.5	.
50	MO	14	7.0	7.0	.
51	WY	8	8.0	0.0	.

control, plus the number of federal aid bridges, with 15,000 (miles plus bridges) being the dividing line between large and small states. This results in 29 states being included in the "large" category and 22 states in the "small" category. Measures of system size were used, rather than area, to better relate the data to fiscal expenditures.

A series of cluster analyses was then run on each of these structures, using Ward's minimum distance cluster analysis. The Ward's model is a straightforward minimum distance model of the following form:

$$D_{KL} = [X_K - X_L]^2/(1/N_K + 1/N_L)$$

where

X = sample mean vector, and
N = number of observations.

In this method, the distance between two clusters is the ANOVA sum of squares between the two clusters, added up over all variables. As observations are grouped into clusters, within-cluster sum of squares is minimized over all partitions obtainable by merging the two clusters from the previous generation. The sums of squares are easier to interpret when divided by the total sum of squares to give proportions of variance. As noted in the SAS procedures manual (*12*), the method tends to join clusters with a small number of observations, and is strongly biased toward producing clusters with roughly the same number of observations. It is also quite sensitive to outliers.

The resulting cluster structure can be viewed as a tree, in which clustering occurs at higher levels as one moves up the tree. Corresponding to these different cluster levels is an associated number of clusters. The choice of a particular number of clusters with which to identify the result is essentially arbi-

TABLE 5 STATES RANKED BY "MISERY INDEX"

OBS	STATE	MISERY	CONDITION	TAXES	MONEY
1	NC	3.26	0.76	0.24	2.25
2	SC	2.46	0.30	0.26	1.90
3	GA	2.44	0.56	0.15	1.73
4	WV	2.21	0.84	0.21	1.15
5	MO	2.16	1.01	0.14	1.01
6	VA	2.06	0.49	0.22	1.35
7	AR	1.88	0.63	0.20	1.04
8	NH	1.67	0.67	0.28	0.72
9	ME	1.65	0.39	0.28	0.97
10	MT	1.64	0.82	0.32	0.50
11	ND	1.61	0.58	0.26	0.77
12	NE	1.59	0.62	0.29	0.66
13	VT	1.59	0.61	0.27	0.70
14	KS	1.57	0.86	0.22	0.49
15	MS	1.52	0.80	0.19	0.53
16	OK	1.48	0.75	0.18	0.55
17	WI	1.46	0.68	0.32	0.45
18	SD	1.46	0.46	0.26	0.73
19	PA	1.45	0.65	0.24	0.56
20	TX	1.42	0.54	0.20	0.68
22	KY	1.40	0.59	0.20	0.60
23	CT	1.39	0.91	0.30	0.17
24	AK	1.38	0.53	0.16	0.68
25	CO	1.35	0.67	0.25	0.43
26	WA	1.32	0.51	0.36	0.44
27	LA	1.31	0.66	0.32	0.32
28	NM	1.27	0.49	0.22	0.55
29	IN	1.25	0.62	0.22	0.41
30	MI	1.24	0.68	0.30	0.26
31	RI	1.23	0.73	0.26	0.24
32	MD	1.23	0.82	0.27	0.13
33	IA	1.21	0.57	0.28	0.35
34	TN	1.21	0.60	0.23	0.37
35	AL	1.20	0.58	0.27	0.35
36	NY	1.17	0.79	0.18	0.20
37	MN	1.07	0.39	0.34	0.34
38	OR	1.06	0.29	0.18	0.59
39	DE	1.05	0.37	0.22	0.45
40	OH	1.02	0.40	0.24	0.37
41	CA	1.02	0.62	0.18	0.21
42	MA	0.97	0.58	0.22	0.16
43	NJ	0.96	0.72	0.16	0.08
44	HI	0.95	0.65	0.17	0.12
45	IL	0.92	0.47	0.26	0.18
46	NV	0.87	0.19	0.24	0.44
47	UT	0.84	0.28	0.28	0.28
48	WY	0.73	0.30	0.08	0.34
49	FL	0.70	0.32	0.19	0.17
50	AZ	0.64	0.14	0.26	0.23
51	DC		1.41	0.31	

Money = 10,000 x Miles of Road/Disbursements
Misery = Condition + Taxes + Money

trary, so for purposes of this analysis we selected approximate breakpoints that clearly identify four, five, or six relatively clean clusters. The selection of this breakpoint was guided by technical data on the clusters produced by the cluster variance program. In particular, we looked for local maxima for Ward's pseudo T-squared, a measure of the strength of the clusters, somewhere in the five-six cluster range.

Figures 2 through 4 show the structure of these clusters by state. A number of general observations are apparent from these results:

- The cluster analysis was in general, not particularly effective. The best cluster structure identified was able to account for only about 25 percent of the variance in the state fiscal data sets. In fact, at the particular cluster level shown (that is, four to six groups) generally less than 10 percent of total variance was accounted for by the cluster. In other words, while these clusters partition the states into separate groups, the within-group variance remains at 90 percent of the total. Clearly, assessments of similarity based on clustered data such as this should recognize that the variations in the data are not producing strongly similar groups.
- Partitioning the data by state size was found to be largely duplicative of the clustering effort itself, since the clusters based on all states tended to show groupings similar to those based on state size. Therefore, further analyses based on size alone are not reported, that variable being subsumed by the full cluster model.
- The number of variables used was a particularly important item in identifying the overall cluster picture. Clusters based on a small number of variables, particularly 7 and 13, showed less strength and greater variation in results. On the

TABLE 6 FACTOR ANALYSIS OF SHA DATA

	Variables	Loadings	Used in Cluster Scheme 19 Vars	12 Vars	7 Vars
Factor 1					
"Congestion,	UPAVC1	.95	x	x	x
system size,	UPAVC2	.95	x		
agency size"	UVC2	.90	x		
	UPATOT	.91			
	UVC1	.85	x	x	
	VMT	.87			
	COMPTR	.89			
	TOTENG	.88			
	RESPOP	.83			
	PLNRS	.87			
	UVCTOT	.80			
	UPSRMI	.80			
Factor 2					
"Expenditure/mile"	DISBPUSC	.94	x	x	x
	CORBPUSC	.93	x	x	
	ADMPUSC	.92			
	MTSPUSC	.90	x		
	PVC	.67			
Factor 3					
"Urban mileage,	UUSC	.78			
rural congestion,	RVC2	.77	x		
maintenance	SERVMTS	.73	x	x	
expenses, fuel tax	FUELTAX	.72	x	x	
revenues"	TOTUSC	.63	x	x	x
	RUSC	.62			
Factor 4					
"Bridges"	NOBRIDGE	.72	x	x	x
	NFANOBR	.72			
	FANOBR	.64			
	TRUM	.72			
	RITOT	.57			
	RPATOT	.59			
Factor 5					
"Deficient Bridges"	DEFBR	.86	x	x	x
	PDFNABR	.80			
	PDFABR	.76			
	FADEF	.63	x		
	NFADEF	.62	x		
Factor 6					
"Tax rates,	DISBPRMV	.67	x	x	
disbursement rates"	DIESLTAX	.67			
	DISBPOP	.65			
	RIPSRMI	.67	x	x	x
	GASTAX	.63	x	x	x
	PSR	.54	x		

TABLE 7 VARIANCE EXPLAINED BY FACTOR MODEL

Factor	Eigenvalues	Percent	Cumulative %
1	23.89	38	38
2	9.01	14	52
3	4.48	7	59
4	4.17	6.6	66
5	3.25	5	71
6	2.48	4	75

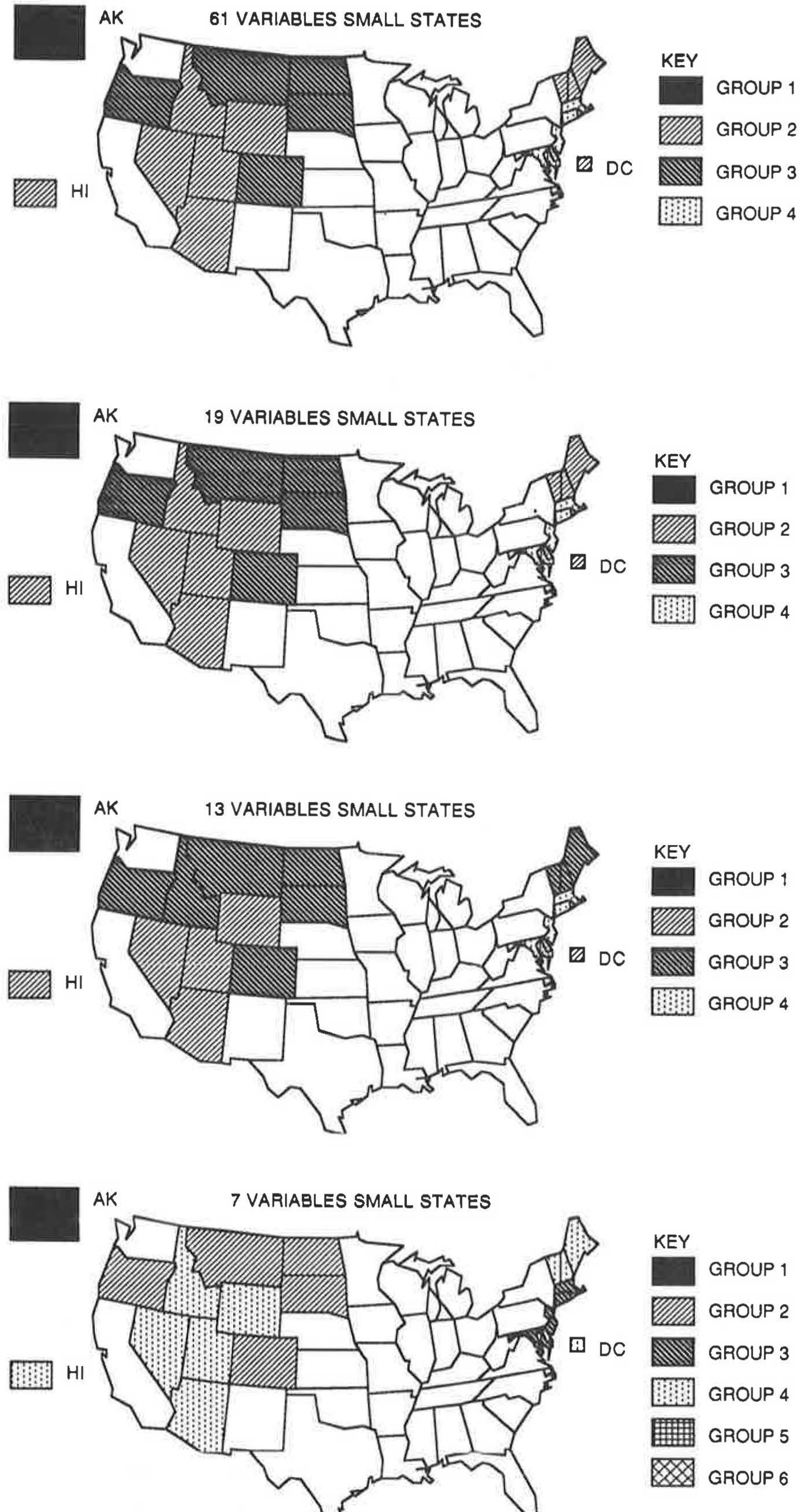

FIGURE 2 Clusters for small states.

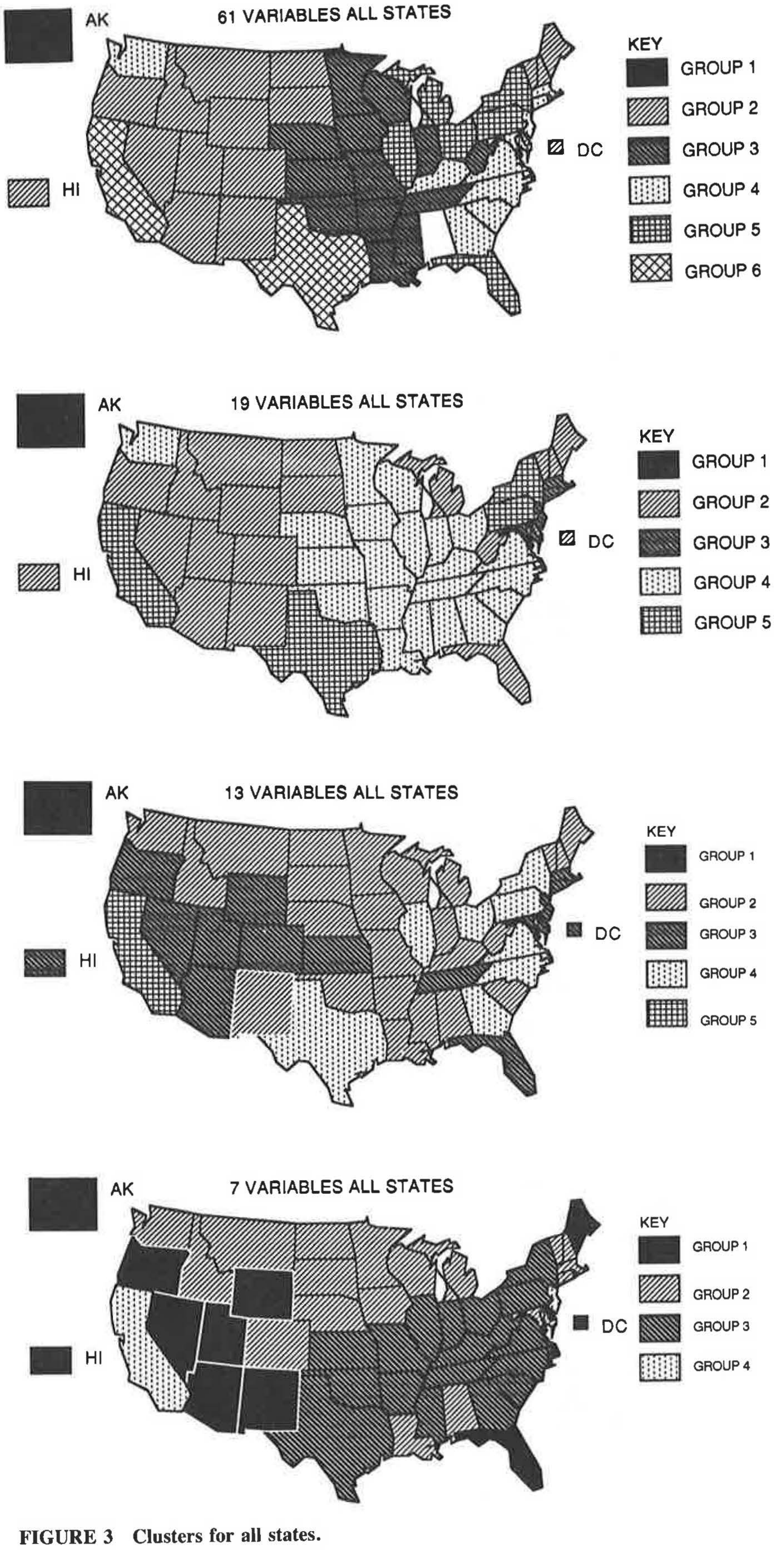

FIGURE 3 Clusters for all states.

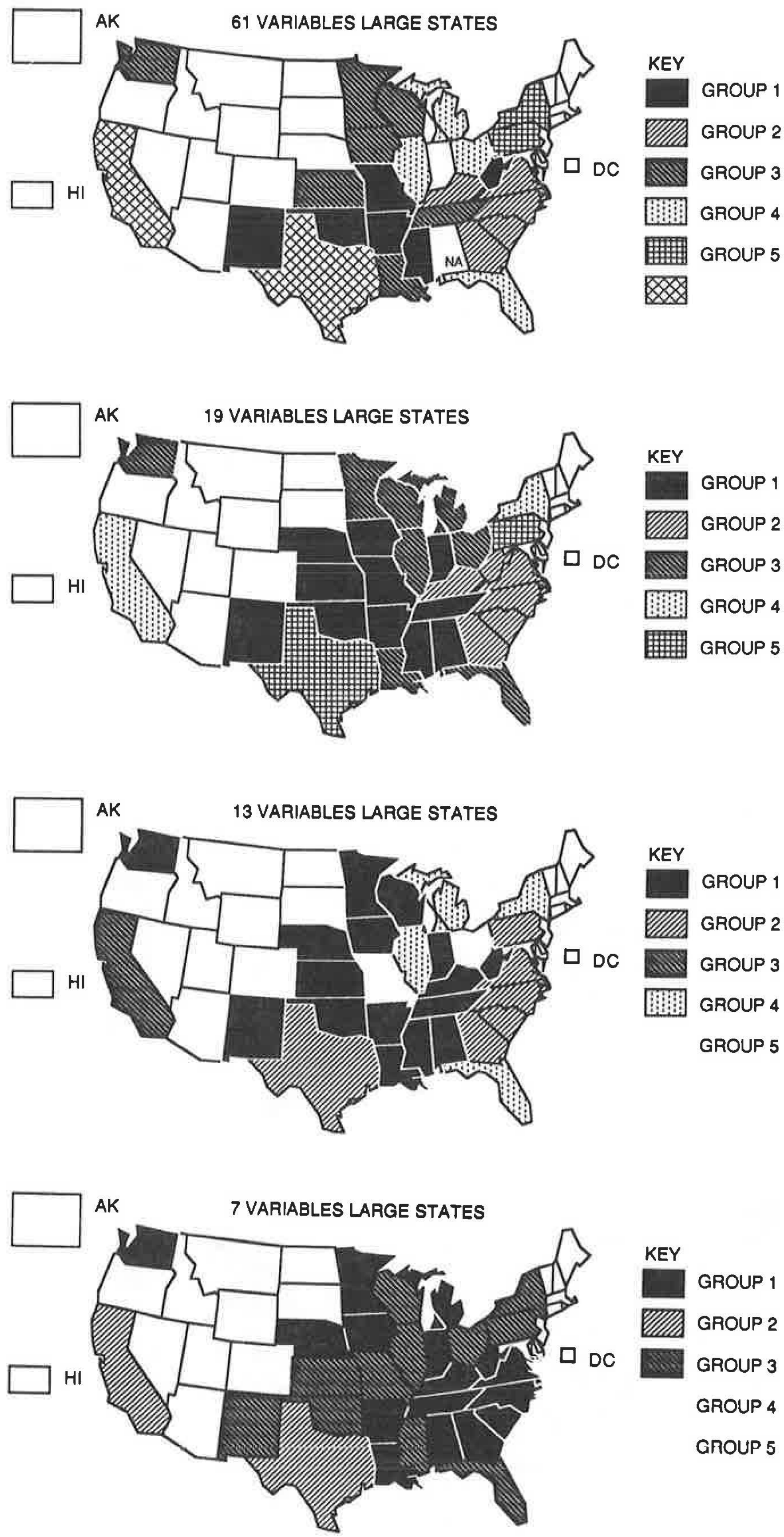

FIGURE 4 Clusters for large states.

other hand, clusters based on a large number of variables (61 variables) tended to produce more cluster groups and greater discrimination, particularly among large states. On balance, it was concluded that clusters based on 19 variables were approximately the right scale for further assessment. In particular, the results now to be discussed relate to those in Figure 3b (19 variables, all states). Figure 5 shows that the tree (dendogram) grouping for this clustering.

● The patterning shown in Figure 3b illustrates both the usefulness and the limitations of the technique. The Figure shows a fairly strong cluster structure in which states are grouped into five general categories:

1. Alaska, by itself.
2. A group consisting primarily of far western and northern New England states, supplemented by West Virginia, Michigan, the District of Columbia, Florida and Hawaii. Important subclassifications of this group are the western block (particularly Colorado, Idaho, and Montana), (North and South Dakota) and the northeast block (particularly Maine, New Hampshire, and Vermont).
3. A mid-Atlantic block of small states, particularly Connecticut, Massachusetts, Maryland, and New Jersey.
4. A large midwestern/southern block consisting of central state subgroups (Alabama, Tennessee, Indiana, Arkansas, Kentucky, and Iowa), southern states (Georgia, North Carolina, South Carolina, and Virginia) and midwest states (Kansas, Missouri, and Oklahoma).
5. Large states (California, New York, Texas, and Pennsylvania).

The dendogram (Figure 5) associated with this cluster shows particularly sharp breaks between Group 5 and other groups, with much less variation and much weaker breakings for the other groups.

While there is considerable overlap in these groups, Table 8 shows some interesting patterns. The mid-Atlantic and big states (Groups 3 and 5) spend two or three times as much per mile as do the other groups, yet their overall condition is not much better. The mid-Atlantic block averages about as much traffic as the midwestern/southern block, and outspends it almost 4:1 per mile, but has about 3 times as many congested miles of road, and only slightly better road and bridge conditions.

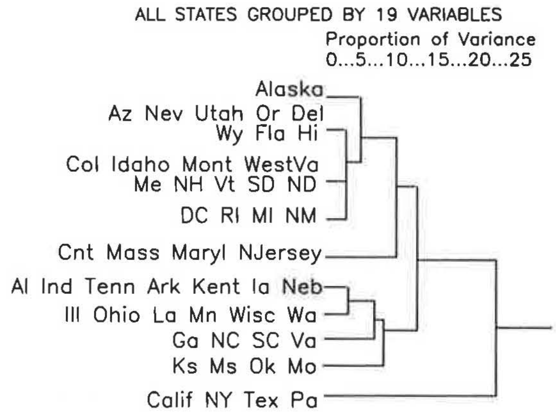

FIGURE 5 All states grouped by 19 variables.

This paper goes no further with this relative performance comparison of the states. That is a topic for later research.

DISCUSSION OF RESULTS

The findings above suggest a number of important policy implications for state decision making and political coalitions. First, the overall weakness of the results suggest that coalitions based on factors relating to current issues and other unaccounted for dimensions are certainly important and need to be nurtured. An analytical tool like this can provide some guidance as to what states may have similar problems, but it is not a substitute for good coalition building.

TABLE 8 COMPARISON OF CLUSTER MEANS

Variables	1 Alaska	2 Far West/NE	3 Mid-Atlantic	4 Midwest/South	5 Big States
Miles under state control	11,426	7,627	3,978	24,752	37,513
Bridges	835	4,152	4,572	17,172	26,358
VMT (million)	3,589	15,830	35,906	35,298	123,959
Gasoline tax, ¢	8	12.3	11.9	11.9	9.75
Fuel tax revenues (000)	$24,197	$76,577	$150,729	$189,386	$471,626
Disbursements/mile	$28,697	$43,320	$222,503	$34,092	$63,767
Capital disb./mile	$14,544	$25,283	$77,065	$20,008	$31,856
Maintenance disb./mile	$8,816	$7,525	$32,346	$6,136	$14,328
Interstate PSR % <2.5	29.6	9.4	4.2	4.2	5.8
Mi. of congested urban Interstate	9.0	19.3	138.7	75.0	·341.0
Mi. of congested urban principal arterial	0.0	10.4	65.5	17.6	295.75
% Deficient bridges	19.9	33.5	41.0	48.0	42.1
Number of states	1	21	4	21	4

Secondly, the dendogram in the cluster analysis suggests that large states, particularly the four largest, are considerably different from the remainder in expenditures and power, as well as size. There is a considerable difference between these four and all the other states, even though the other states also contain numerous regional groupings. In looking at the differences between states, one should recognize that the other 47 are far more similar to each other than they are to these four.

Third, a number of interesting regional clusters appear that might have been expected. These include: western, southeastern, northern New England, and southern New England blocks. Within the dendogram one can also see strong regional affiliations in the Midwest as well. In a sense, therefore, the analytical tools here confirm the importance of geography in producing similar fiscal patterns. It should be noted that no geographic closeness measure was included in these statistics, and therefore the clustering of the state groups by geography is a testament to the importance of geographic similarity in producing fiscal similarity.

On the other hand, certain liaisons appear that might not have been anticipated. These include the northern-tier liaison between Maine, New Hampshire, and Vermont on the east and North and South Dakota on the west; the mountain liaison between Colorado, Idaho, and Montana on the west, and West Virginia in the east; the recreational liaison between Florida and Hawaii; and the western group consisting of Arizona, Nevada, Utah, Oregon, and Wyoming. Other important subregional groups are the southeastern seaboard core states of Georgia, North Carolina, South Carolina, and Virginia; a far-flung quadrangle consisting of Louisiana, Minnesota, Wisconsin, and Washington; a midwestern axis of Illinois and Ohio; and a broad southern midwestern axis of Alabama, Tennessee, Indiana, Arkansas, Kentucky, Iowa, and Nebraska.

The analysis raises more questions than it answers. Why are geographic liaisons so important in classifying agency expenditure patterns? What leads to far-flung similarities in apparently isolated environments? How can the similarity structures here be leveraged to identify and form working political coalitions? Would results be different if data were available over time, if data were more recent, or if additional information relating to modes of travel other than highways were included? Does the cluster methodology itself produce results which are different? What do the results say about the relative performance effectiveness or efficiency of the State highway agencies? To what extent can the effects of weather and climate be interpreted as influencing these results? Answers to these and numerous other questions must weigh the results of later research.

ACKNOWLEDGMENTS

The author acknowledges the assistance of Nathan S. Erlbaum of the New York State Department of Transportation for initial development of this data set. Ralph Campano and Steve Sonderman of UNC-Charlotte assisted in conversion and analysis. Norma Redmond painstakingly prepared the manuscript. The author of course retains the responsibility for errors of fact or omission.

REFERENCES

1. C. D. Phelps. Real and Monetary Disbursements of State and Local Highway Investment. *American Economic Review*, Vol. 59, No. 4, Sept. 1969, pp. 507–521.
2. K. Rao. *NCHRP Synthesis of Highway Practice 100: Managing State Highway Finance*. TRB, National Research Council, Washington, D.C., 1982.
3. H. R. Trumbower. "Collection and Disposition of Motor Vehicle Revenues," HRB, Proceedings, 1927.
4. *Highway Statistics, 1984*. FHWA, U.S. Department of Transportation, Washington, D.C., 1985.
5. *Highway Statistics: Summary to 1985*. FHWA, U.S. Department of Transportation, Washington, D.C., 1987.
6. C. W. Woodward. *Trends in Federal Domestic Transportation Programs, Revenues and Expenditures by State, Fiscal Years 1957-75*. U.S. Department of Transportation, Washington, D.C., 1979.
7. *Highway Taxes and Fees*. FHWA, U.S. Department of Transportation, Washington, D.C., 1984.
8. *Financing Federal-Aid Highways*. FHWA, U.S. Department of Transportation, Washington, D.C., 1987.
9. *Special Report 207: Transportation Professionals*. TRB, National Research Council, Washington, D.C., 1985.
10. *Sixth Annual Bridge Report to Congress*. FHWA, U.S. Department of Transportation, Washington, D.C., 1987.
11. K. Obeng. Application of Path Analysis Methodology to Transit System Maintenance Performance. *Transportation*, Vol. 15, No. 4, 1989, pp. 297–316.
12. *Statistics: User's Guide*. SAS Institute, Cary, N.C., 1985.

Publication of this paper sponsored by Committee on Statewide Multimodal Transportation Planning.