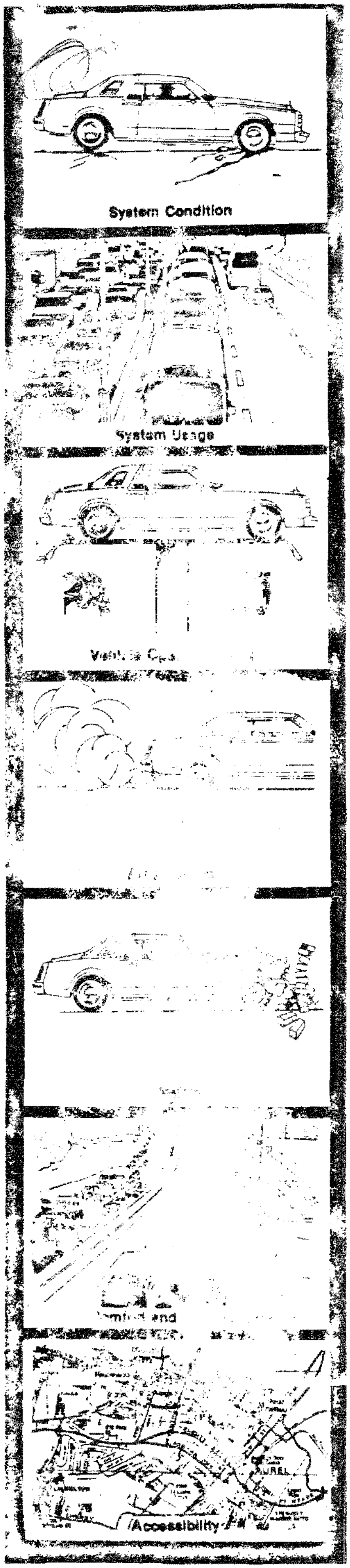


HIGHWAY PERFORMANCE MONITORING SYSTEM

TECHNICAL REPORT NUMBER 3

A STATISTICAL ANALYSIS

U. S. DEPARTMENT OF TRANSPORTATION
Federal Highway Administration
Program Management Division

October 1978

Highway Performance Monitoring System

A Statistical Analysis

by Anthony P. Kancler

TABLE OF CONTENTS

LIST OF TABLES AND ILLUSTRATIONS

LIST OF APPENDICES

Background

The HPMS effort is directed to the design and development of a continuing frame of reference which is capable of assessing the performance of highway systems with respect to the safe, efficient, and economical movement of people and goods, and which is also a basis for forecasting the potential impacts of future alternative programs and policies on such performance.

The development of a framework for monitoring changes in the physical and operational characteristics of highways requires the identification of specific data elements necessary for the derivation of performance measures pertinent to national policy planning and program evaluation concepts. Preliminary to this paper, there was developed a tentative group of data elements considered essential, in varying degree, to the measurement of highway performance.[1/] These specific data elements were selected with the underlying purpose of minimizing ongoing and unique data requirements.

From the beginning of the conceptual development of HPMS, it has been envisioned that HPMS data will be gathered using statistical sampling techniques. This paper is concerned with the efficacy of proposed sampling methods in obtaining essential data element values. While it is assumed there is a "technically best" way to collect sample data element estimates, manpower and cost considerations are necessary to arrive at a usable sampling plan. The selected sampling plan for monitoring highway performance must be simple and yet cost efficient, be applicable to individual State needs, be suitable for data aggregation by functional system at regional and national levels, be capable of detecting statistically significant changes over time, and since time phasing is a major concern for State workloads, the plan must be adaptable to the resampling requirements specific to each data element.

The proposed sampling plan having the above requisites is a randomly selected primary sample or "panel" of road sections, generally of unequal length, which will remain fixed after the initial sample selection process. The primary sample and tentative subsample data from this fixed panel of road sections can be inventoried on a cyclical basis. The advantages of this plan are outlined in the Sample Design section of this paper.

In addition to the primary sample and subsample for HPMS data collection requirements, there are two other data collection categories not covered in this paper; namely, case studies--limited to necessary typical data that cannot be obtained from a sample of sections, and also areawide nonsection specific data.

No consideration has been given to the sampling of local roads. The consensus is that these roads are not as sensitive to changes in the highway performance elements as are the five higher level functional

---

1/Appendix A--Primary Data Elements

road categories. If they are to be sampled, the recommendation is to sample the relatively high-volume local roads on a road section basis with volume group stratification. The remaining low-volume roads may be sampled on an area basis according to the method for estimating local rural and urban VMT in Technical Report 31, July 1973, Sampling Surveys for Estimating Local, Rural, and Urban Vehicle Miles of Travel, U.S. Department of Transportation, Federal Highway Administration.

Assuming that the concept of the fixed panel is reasonable, this paper will address these issues:

(1) Is the use of nonuniform length road sections as sampling units acceptable for a predetermined level of accuracy?

(2) What data elements are critical to sample size?

(3) What number of randomly selected road sections will produce reliable statistical estimates, conforming to predetermined precision level requirements, for a sufficient number of data elements influencing highway performance?

The sections that follow are concerned with a fairly detailed description of road section sampling applications, sample design and size requirements, and an analysis of test results. Reference is made to the General Assessment section of this paper for the reader who desires a general overview of the qualifications and limitations of the proposed fixed panel sampling plan.

Road Section Application

One of the first concerns in the planning of a sample design for estimating data element values needed in the monitoring of highway performance is the acceptability of nonuniform road section lengths as primary sampling units. Traditionally, most formulas used for the sampling of various highway statistics rely upon the uniform mile as the basic sampling unit. Conversely, field data records in the States are for the most part preferably collected from and based on nonuniform section lengths; consequently, the sampling of nonuniform length road sections is the more practical from the standpoint of availability and convenience. A general definition for a nonuniform road section is a measured length of road delimited by any of the following: Intersecting roads or streets, bridges, railroad crossings, natural or manmade barriers, demarkations of convenience, or any combination of the preceding.

The sampling of nonuniform sections for the production of reliable quantitative data element estimates required section-length weighting adjustments; data values in uniform 1-mile sections are intrinsically self-weighting. The 1976 National Highway Inventory and Performance Study (NHIPS) tape stored data provided a convenient source for testing whether samples of selected data elements using 1-mile sampling units compared with similar samples of nonuniform section-length sampling units

produce statistically significant different estimates. In the application of tests to selected States, it was found, as expected, that a comparison between the two groups showed statistically significant differences in quantitative data estimates when the estimates for the nonuniform length sections excluded the weighting adjustment. The use of appropriate formulas featuring section-length weighting reduced these differences considerably.

In contrast to quantitative data, the estimates for qualitative values (proportions) generally require no weighting adjustments. The proportions for selected attributes of data elements (for example, the percentage of pavement condition as poor, fair, or good) are expressed as ratios of sampled attribute mileage to total sampled mileage, regardless of whether the length of the sampled sections are uniform or nonuniform.

The number of nonuniform length sections* available for sampling in rural areas tends to be less than those for uniform 1-mile units in rural areas because the median section length exceeds 1 mile; the reverse is true for urban areas. It should be noted that this relationship does not necessarily have bearing on sample size requirements. Table 1 is derived from the nationwide NHIPS study file and gives median values for section lengths by area and functional class.

Table 1

Rural	Median Section Length	Urban	Median Section Length
Interstate	3.9	Interstate	1.0
Other Principal Arterials	2.0	Other Freeway	0.9
Minor Arterials	1.8	Other Principal Arterials	0.4
Major Collectors	1.8	Minor Arterials	0.4
Minor Collectors	1.6	Collectors	0.3

Sample Design and Time-Phasing Considerations

In order to maintain sample design simplicity, consideration is given to two simple random sample methods:

(1) A fixed number of sample sections is randomly selected within each functional highway system in each State area (rural, urbanized, or small urban). This method is rejected because the heterogeneous mix of data element values require an exceptionally large number of samples to achieve desired accuracies.

(2) The fixed number of sample sections is randomly selected within predetermined Average Daily Traffic (ADT) groups within a given functional highway system within State area. This is the recommended sample design. The

*Further reference to "sections" will be assumed to mean nonuniform length sections.

allocation of sampling units into relatively homogeneous ADT volume groups produces estimates of greater accuracy with respect to ADT for a smaller number of samples at functional class summation levels. Stratification by ADT volume has other advantages: (1) It serves as a weighting device for quantitative data element values sampled from sections of nonuniform length; (2)the effect of volume on volume sensitive data element values may be measured; and (3) it is useful in the application of specific statistical formulas, such as ratio estimates for Daily Vehicle Miles Traveled (DVMT) estimates.

Once the sampling structure is established, the frequency of sampling operations must be taken into account. One way is to select a new sample of roadway sections from which primary sample data are to be inventoried on a cyclical basis. This method for obtaining data is not recommended for two major reasons: (1) A new sample must be taken each cycle with attendant instruction and coordination problems; and (2) even though the sample design uses statistical concepts to the maximum extent possible, changing samples may introduce sampling errors larger than the changes in performance over a period of time, thus destroying the comparability of performance and impact measures over time. ?

The recommended way is to allow the sample of road sections or "panel" to remain "fixed" after the initial selection. The primary sample can then be inventoried from this fixed panel on a cyclical basis for the desired data element values. The advantages of this method are: (1) The need for the periodic drawing of an entirely new sample is eliminated, thereby keeping the sampling error relatively constant and increasing the validity of comparisons over time; (2) the need to sample many of the data elements frequently is eliminated because many data element value estimates remain unchanged until a capital improvement is made, and such an improvement can be used as a signal to initiate an update for the affected section; (3) the data elements that consistently change in time can be updated on a cyclical basis, the cycle interval being dictated by the characteristics of the individual elements, the intended use, and the time/cost considerations; and (4) for the first time, a statistically sampled "fixed panel" of sections will be established that can be a great aid in yet undefined special studies.

The concept of the "fixed panel" allows for the making of minor adjustments for changes in urban boundaries. It assures statistically valid comparisons of performance measures over time for impact assessments and should contribute to long-range economies of effort.

Should include disadvantages of panel.

Sample Size Requirements

In addition to the sample design and data collection time-phasing considerations of the HPMS, sample size requirements for a desired level

of accuracy for data element estimates must be determined. Two basic methods have been considered for obtaining the number of road sections to be used as primary sampling units: (1) The required number of samples is determined by the application of specified sampling rates to the total number of available sampling units in an individual ADT volume group. This use of sampling rates is not acceptable because there is no assurance that desired levels of accuracy will be attained. Observations from NHIPS test data show that for a given accuracy level the required sampling rate may be from 2 percent to 100 percent, depending on the variability of the data element values and the size of the given volume group. Also, if different sampling rates are used for different data elements, interrelationships which are essential in the determination of performance measures, e.g., ADT and pavement condition, will not be known. (2) The required number of samples is derived empirically by formula from the normal dispersion characteristics of ADT values within the recommended framework of preselected ADT volume groups or strata. The advantage of this method is its implicity and general applicability. There is no need for obtaining necessary data element variances from pilot studies in the field. The sample size requirements obtained by this method relate to the critical data element ADT, whose values can be conveniently stratified to advantage. Sample sizes for desired levels of accuracy can be computed in the Washington Office with the aid of individual State road section information available in the 1976 NHIPS file. Computed sample size requirements for each volume group within the five functional systems can be assigned to the individual States for random allocation, preferably computerized, in accordance with the State's highway network characteristics.

The formulas for the computation of sample size by the recommended empirical method are presented and illustrated in Appendix B.

Sample size requirements will vary by State according to the number of available road sections and the statistical parameters of the predetermined ADT volume groups. The application of the empirical method to NHIPS data for a group of States was helpful in obtaining an expected range of State areawide sample size requirements for selected precision levels. The term "precision level" in this report is defined as the maximum allowable error in a sample estimate at a given confidence level; or expressed conversely, the degree of confidence that the error of a produced estimate will fall within a desired fixed range. Thus, for a precision level of 80 percent confidence in a 10 percent allowable error, there is a probability of 80 times out of 100 that the error of a data element estimate will be no greater than 10 percent of its true value. For the basis of this evaluation, it is tentatively recommended that the HPMS sample size requirements be based on a precision level of 80 percent confidence in an allowable error of 10 percent, viz 80-10. A generalized range of areawide sample size requirements applicable to all States for selected precision levels is shown in Table 2.

SAMPLING

Table 2

Confidence Level and Percent Allowable Error	Range of Sample Size Requirements by Area Within State - Number of Sections for Functional Systems		
	Rural	Individual Urbanized Areas	Small Urban
70-10	100-400	50-375	30-325
80-10	135-550	70-500	40-425
95-10	270-1100	140-1000	80-850
80-5	400-1650	210-1500	120-1275
95-5	800-3300	420-3000	240-2550

General Assessment

Tests conducted on the 1976 National Highway Inventory and Performance Study (NHIPS) for three States show as a whole that a fixed sample panel of road sections selected by random sampling within Average Daily Traffic (ADT) volume strata gives data element estimates conforming to desired accuracy requirements for quantitative values. Estimates for qualitative or proportionate values obtained from the same sampling bases are also acceptable but with some limitations. As with most field collected data, the quality of the final output depends on the accuracy of the source material. Poor information furnished by a State will result in incorrect sampling errors. This has occurred in a few instances in this study on test data.

Sample size requirements for the predetermined level of 80 percent confidence in allowable errors of estimate no greater than 10 percent of the true value (100 percent sampling) range from 8 to 122 road sections for the three States tested, with an average of 65 road sections per area functional system. It is felt that the sample size requirements for an 80-10 precision level will not be overly demanding and will ensure a reasonable degree of accuracy in data element estimates. The ranges for the examined three States for areawide sample size requirements for all functional systems, excluding local roads, are: Rural 260-460; Urbanized 225-400; and Small Urban 90-290.

The use of an empirical method for determining sample size requirements in the Washington Office for all States, along with the availability of the NHIPS data file, is a distinct advantage. The random allocation of the selected samples to the sample panel and also the data collection is left to the individual States.

Although only 8 of a total of 33 data elements in the highway supply and use data categories in the three States were tested, there was sufficient variety in the tested element characteristics to assume reasonably that reliable estimates for quantitative statistics such as data element means or aggregate totals can be produced for most data elements at a predetermined level of accuracy. This assumption holds

no mention of aggregation to statewide values

for the five functional systems in most if not all States. Also, it appears that the use of ADT volume group stratification for the selection of primary sampling units is not incompatible with the characteristics of any specific data element.

The presence of variable length in the sampled road sections is offset by road length weighting adjustment factors included in the formulas for estimating quantitative data element statistics. For qualitative data, the proportions for specific data element attributes are obtained by the simple ratio of sampled attribute mileage to total samples mileage in a functional system.

In contrast to the overall acceptable test results on quantitative data, the application of a precision level of 80-10 to qualitative (proportions) data is limited. The use of relatively high precision levels as standards of accuracy for the estimated proportionate values of attributes or classes within selected highway data elements requires a far greater sample size than that needed for quantitative values. Even though the preselected precision level of 80-10 is generally applicable to all quantitative data in this sample panel, the average sample size in this study shows that it can apply only to proportionate values of approximately 70 percent or higher. Since the level of accuracy for estimated proportions is closely related to sample size, this proportion can be significantly lowered only by larger increases in sample size.

Correcting the imbalance in sample size requirements for quantitative and qualitative estimates can be best described as a numbers game trade-off--large increases in sample size will result in an "overkill" in the accuracy requirements for quantitative estimates and a modest improvement in the proportion estimates; moderate increases in sample size will produce some upgrading in the precision level of quantitative estimates but will be insignificant for proportions.

The recommendation is to maintain the sample size level of this study for practical reasons. Assuming that the desired accuracy for proportion estimates is limited, the estimates so produced nevertheless serve as fixed panel benchmarks for monitoring change at prescribed time periods. The study sample is capable of detecting statistically significant changes of 10 percent or greater in proportion estimates for areawide functional systems during the monitoring process of a State. Based on the average sample size of this study panel, absolute changes of less than approximately 10 percent in the estimated value of a proportion would not be considered statistically significant because of the wide margin of sampling error. Detecting statistically significant changes of 5 percent requires the quadrupling of the average sample size of 65 per functional system, regardless of the sample design.

From the standpoint of sample size and cost constraints, we must use a functional system sample size capable of detecting changes of at least 10 percent, and preferably less, in proportions. Such a sample size

will ensure an accuracy standard of at least (80-10) in the measurement of quantitative values at the functional system level. Moreover, the improvement in accuracy obtained by aggregating functional system estimates to higher geographic levels (e.g., individual urbanized area aggregates) can be considerable, especially in estimates of proportions.

Further details concerning the content of this section are presented in the next section, the Analysis of Test Results.

Analysis of Test Results on NHIPS Data

The goal of precision tests conducted on the 1976 National Highway Inventory and Performance Study (NHIPS) data is to ascertain, (1) whether a specific, fixed sample design of selected road sections will produce reliable statistical estimates for a sufficient number of data elements influencing highway performance; and also, (2) whether these statistical estimates conform to predetermined precision level requirements. The form of sample design for these tests was a simple random sample stratified by fixed ADT volume groups within functional highway class within type of area (rural, urbanized, or small urban).

As mentioned in the section on Sample Design and outlined in Appendix B, the critical element Average Daily Traffic (ADT) was selected as the data element on which to base sample size requirements--i.e., the required number of road sections (primary sampling units) per ADT volume group for a predetermined precision level of 80 percent confidence in an allowable error of 10 percent. This sampling base was used to produce quantitative and qualitative estimates of selected data elements; namely, daily vehicle miles traveled, pavement condition, percent trucks, K-factor, V/C ratio, access control, lane width, and right shoulder width. Tests were limited to the NHIPS data for three States--Arizona, California, and Pennsylvania.

After the sample size for preselected volume groups within functional system for each State had been computed, a computerized random selection program was used to select the needed number of road sections from the NHIPS file for each volume group. The road sections so selected served as the fixed "panel" of primary sampling units.

Table 3 shows sample size by the five functional classes within area for each of the three surveyed States. For the sake of brevity, sample sizes by volume group within functional classes are not itemized. Also shown are the area totals for the number of road sections available for sampling in each State and the overall sampling rates.

80-10 level

Table 3

STATEWIDE SAMPLE SIZE BY FUNCTIONAL CLASS, TYPE OF AREA, AND STATE

RURAL

	Interstate	Other Principal Arterials	Minor Arterials	Major Collectors	Minor Collectors	Area Sample Size	Total Sections in State	Percent Sampled
AZ	36	53	72	60	40	261	1960	13.3
CA	39	89	109	108	115	460	9525	4.8
PA	30	122	91	89	113	445	14181	3.1

URBANIZED

	Interstate	Other Freeway	Other Principal Arterials	Minor Arterials	Collectors	Area Sample Size	Total Sections in State	Percent Sampled
AZ	21	24	101	32	47	225	2149	10.5
CA	50	64	70	113	104	401	21814	1.8
PA	52	68	91	69	95	375	6749	5.6

SMALL URBAN

	Interstate	Other Freeway	Other Principal Arterials	Minor Arterials	Collectors	Area Sample Size	Total Sections in State	Percent Sampled
AZ	9	—	39	24	21	93	829	11.2
CA	22	25	69	79	93	289	4236	6.8
PA	8	29	69	56	64	226	3569	6.3

TOTAL - ALL AREAS

						Total Sample Size	Total Sections in State	Percent Sampled
AZ	—	—	—	—	—	579	4938	11.7
CA	—	—	—	—	—	1150	35575	3.2
PA	—	—	—	—	—	1046	24499	4.3
Three States	—	—	—	—	—	2775	65012	4.3

Estimates of the statistical parameters for the quantitative and qualitative values of the eight data elements were computed for the functional system levels. Where applicable, these were the aggregate totals, the means and average proportions, the standard errors, and the relative errors of the estimated means or proportions. The formulas used in calculating these estimates are not shown in this report but are available as a package applicable to most data elements involved in highway performance monitoring.

The statistic, "percent relative error of the estimated mean or proportion," is useful in deciding whether the size of the ADT based panel of primary sampling units is adequate for producing acceptable estimates for all eight data elements in this study at the preselected precision level of 80 percent confidence in a maximum allowable error of 10 percent. The relative error, also known as the coefficient of variation, can be used to set up a precision index expressible in terms of any desired confidence level and allowable error. The degree of confidence that one may expect in the accuracy of a produced statistic may be estimated from the ratio $d/\%R{\cdot}E = Z$.

Where,

d = The allowable percent error of the sample estimate from the true value of a given data element or its attribute.

%R·E = The percent relative error of a produced statistic. For quantities: $\%R{\cdot}E = s_{\bar{x}}/\bar{x}$

where, $s_{\bar{x}}$ is the standard error of the estimated mean of a data element, and $\bar{x}$ is the estimated mean. At 68 percent confidence, the percent relative error is also called the coefficient of variation.

For proportions: $\%R{\cdot}E = s_{\bar{p}}/\bar{p}$ where, $s_{\bar{p}}$ is the standard error of the estimated proportion and $\bar{p}$ is the estimated proportion.

Z = A ratio which when translated into standard error units of area under the Normal Curve gives the confidence level or probability that a produced estimate will not exceed an allowable range of error. The Z-values for confidence levels of 50, 60, 68, 70, 80, 90, and 95 percent are 0.68, 0.84, 1.00, 1.04, 1.28, 1.65, and 1.96, respectively.

In using the above relationship, one finds that a precision level of 80-10 is approximately the same as 50-5[1/]. The data shown in Table 4 are the

1/For a precision level of 80-10, the percent relative error of an estimate must be no greater than 7.8 percent; e.g., $\%R{\cdot}E = d/z = \frac{10.0}{1.28} = 7.8$.

At a 50-5 precision level, the value of Z for a relative error of 7.8 is 5.0/7.8= 0.641, which is a confidence level of 48 percent, roughly 50 percent.

confidence levels for a 5 percent allowable error for the selected quantitative data element estimates. The 5 percent allowable error base is used in order to better illustrate precision deviations from the required minimum of 50-5. A confidence level of less than 50 percent in Table 4 indicates that the accuracy of the estimates derived from the samples data for a given data element does not conform to the predetermined minimum precision requirements.

In Table 4, the row "all functions" class is an indicator of the "average" precision level for the five functional systems within a given area type. The "all functions" confidence level takes into consideration both the variances of respective data elements within each functional class and also the variance between functional classes. For any functional class, it is expected that the dispersion of the variables in different random sample sets will not be quite the same due to random chance variations, road geometrics, and indeterminate manmade causes. Because of these variations, we can only assume in Table 4 that the specified confidence level of 50 percent in Table 4 is the lowest level at which the predetermined 80-10 minimum precision requirements are maintained.

The following equivalency table may be helpful in relating the confidence level values in Table 4 to an 80 percent confidence--allowable error base:

Equivalency Table

Table 4 Confidence Level for 5 Percent Allowable Error	*Percent Allowable Error at 80 Percent Confidence
80	5.0
70	6.0
60	7.5
50	10.0
40	12.0
30	17.5
20	25.0
15	33.5

*Percent allowable error at 80 percent confidence = $5/Z \cdot (1.28)$ where,

5 = Percent allowable error in Table 4
Z = Value of Z for the confidence level shown in Table 4
1.28 = Value of Z for 80 percent confidence in the allowable error

In general, the confidence levels presented in Table 4 for the mean value estimates of the data elements under consideration are satisfactory on an area basis for the three sampled States. By functional class, there are scattered instances where the confidence levels of the estimates are considerably below the desired level of 50 percent. Some of the causes of these low values can be explained; other causes are indeterminate

TABLE 4
Confidence Levels for 5 percent Allowable Error for Quantitative Data of Selected Data Elements

State, Type of Area, and Functional Class	DVMT	Pavement Condition	K-Factor	Percent Trucks	Right Shoulder Width	Access Control	Lane Width	V/C Ratio
ARIZONA								
Rural								
Interstate	58	46	99	42	99		99	62
OPA	70	63	87	76	70		99	39
Arterial	80	63	92	37	38		99	63
Major Collector	70	40	87	46	32		92	43
Minor Collector	78	24	99	37	16		99	45
All Functions	85	61	99	70	54		99	70
Urbanized						Not Applicable		
Interstate	82	66	99	70	99		99	78
OFY	97	99	99	99	99		99	92
OPA	99	85	99	82	44		69	68
Minor Arterial	97	42	99	78	31		50	30
Collectors	88	76	99	99	33		71	42
All Functions	98	95	99	99	54		88	78
Small Urban								
Interstate	91	78	99	84	99		99	99
OFY	none	none	none	none	none		none	none
OPA	80	53	94	51	59		56	51
Minor Arterial	88	82	66	82	21		43	26
Collectors	81	41	99	99	25		70	24
All Functions	84	68	99	94	65		81	55

(More)

Table 4 (Continued)

State, Type of Area, and Functional Class	DVMT	Pavement Condition	K-Factor	Percent Trucks	Right Shoulder Width	Access Control	Lane Width	V/C Ratio
<u>CALIFORNIA</u>						Not Applicable		
<u>Rural</u>								
Interstate	95	56	99	34	99		99	53
OPA	99	90	96	38	76		99	54
Arterial	97	97	93	48	51		99	52
Major Collector	60	39	82	38	49		99	43
Minor Collector	97	85	98	30	46		99	46
All Functions	91	99	99	53	78		99	77
<u>Urbanized</u>								
Interstate	99	98	99	25	99		99	71
OFY	99	99	99	39	96		99	78
OPA	79	95	74	24	69		62	39
Minor Arterial	97	71	61	34	64		63	39
Collectors	68	59	95	34	67		90	33
All Functions	99	91	95	51	91		92	61
<u>Small Urban</u>								
Interstate	92	89	99	32	93		99	85
OFY	87	96	99	45	90		99	52
OPA	99	87	86	39	87		60	66
Minor Arterial	56	71	94	31	62		55	20
Collectors	79	78	97	39	47		65	49
All Functions	83	94	99	56	78		82	41

(More)

Table 4 (Continued)

State, Type of Area, and Functional Class	DVMT	Pavement Condition	K-Factor	Percent Trucks	Right Shoulder Width	Access Control	Lane Width	V/C Ratio
PENNSYLVANIA								
Rural								
Interstate	82	85	66	32	99		99	44
OPA	79	97	99	94	69		99	32
Arterial	63	88	97	94	70		99	33
Major Collector	67	85	93	46	59		97	71
Minor Collector	98	84	99	76	80		88	77
All Functions	91	99	99	93	94		99	62
Urbanized								
Interstate	99	95	98	65	99		99	80
OFY	99	41	99	96	99		92	34
OPA	99	80	99	99	39		25	41
Minor Arterial	85	69	99	99	34		60	71
Collectors	81	74	99	62	72		65	45
All Functions	99	97	99	99	69		55	72
Small Urban								
Interstate	44	86	99	33	99		99	53
OFY	91	98	82	98	99		99	65
OPA	96	84	99	99	45		76	58
Minor Arterial	87	79	98	96	44		70	45
Collectors	99	85	99	69	30		52	66
All Functions	99	98	99	99	66		90	77

or are attributable to chance variations. Average pavement condition rating estimates are influenced by the proportion of gravel and graded roads in a sample which in effect reduces sample size and enlarges the error of estimate. The average right shoulder width estimates are similarly affected by sample sections having no shoulders, especially on collectors and minor arterials. Low confidence levels also occur for some estimated lane width averages because of the mix of lane width and curb to curb measures.

Tables 5 through 10 are functional class within area exhibits of attribute proportions for six of the sampled data elements. The elements DVMT and percent trucks are not included because partitioning by proportion is not justified in this report. Each cell in the tables has the attribute proportions for Arizona, California, and Pennsylvania, respectively. The number of attribute proportions is more or less arbitrarily selected and ranges from three for Access Control to eight for Lane Width. The estimated proportion for each attribute is the ratio of the attribute sampled mileage to the total sampled mileage in a given functional system times 100.

Some general comments on the nature of the distribution of attribute proportions for the selected data elements in Tables 5-10 follow:

Access Control - As expected, the Interstate functional system for all areas and also the "other freeways" system in urban areas have almost completely full or partial access control. The other functional systems are for the most part without access control.

V/C Ratio - For rural areas most of the ratio values are below 0.41. In urban areas, although a major part of the ratio values are less than 0.26, there is a scatter of proportions of varying magnitudes throughout all the value classes.

K-Factor - The central tendency for the highest proportions is in the midvalue class 10-14 for all areas.

Pavement Condition - The highest proportions tend to be generally in the 3.0-3.9 rating class for all areas. The 0.0 rating class is for gravel and graded roads and contains some relatively high proportions for the lower volume roads in Arizona.

Right Shoulder Width - Almost all of the shoulder widths on the Interstate system in all areas exceed 8 feet. For the other systems in all areas, the sizes of the proportions and shoulder widths are randomly distributed.

Lane Width - The design lane width for Interstate roads is 12 feet. There appears to be a central tendency for 12-foot lane widths for the other functional systems in all areas. The proportions in urban areas for widths exceeding 15 feet are generally curb to curb measurements.

Table 5

Access Control: Summary of Estimated Proportions for Selected Data Element Categories*

Area Type and Functional System	Full Access Control	Partial Access Control	No Access Control
Rural			
Interstate	100-93-100	0-7-0	0-0-0
OPA	2-24-24	0-21-5	98-55-71
Arterials	0-0-0	0-13-0	100-87-100
Major Collectors	0-1-0	0-1-6	100-98-94
Minor Collectors	0-0-5	0-0-0	100-100-95
Urbanized			
Interstate	100-97-100	0-0-0	0-3-0
OFY	100-77-86	0-23-9	0-0-5
OPA	0-0-2	11-0-17	89-100-81
Minor Arterials	0-0-0	0-4-0	100-96-100
Collectors	0-0-0	0-0-0	100-100-100
Small Urban			
Interstate	100-100-100	0-0-0	0-0-0
OFY	--54-97	--46-1	--0-2
OPA	0-0-0	1-0-23	99-100-77
Minor Arterials	0-0-0	0-0-15	100-100-85
Collectors	0-0-0	0-0-1	100-100-99

*The proportions for each of the three surveyed States are shown in each cell.

Table 6

V/C Ratio: Summary of Estimated Proportions for Selected Data Element Values *

Area Type and Functional System	.26	.26 - .40	.41-.55	.56-.70	.71-.85	.85
Rural						
Interstate	78-58-60	18-29-18	3-9-11	1-4-11	0-0-0	0-0-0
OPA	68-39-34	26-33-31	1-12-20	2-7-11	2-4-0	1-5-4
Arterials	81-75-64	11-12-20	6-5-10	2-1-0	0-2-0	0-5-6
Major Collectors	89-78-68	5-10-14	0-7-8	5-4-6	0-0-4	1-1-0
Minor Collectors	99-87-84	1-5-13	0-3-1	0-0-2	0-0-0	0-5-0
Urbanized						
Interstate	5-33-3	30-19-19	26-7-9	14-13-23	12-7-18	13-21-18
OFY	39-17-37	27-15-35	12-26-10	12-2-1	10-19-1	0-21-16
OPA	33-58-31	9-3-8	8-2-11	9-3-9	14-4-2	27-30-39
Minor Arterials	20-59-58	18-6-10	16-6-12	1-9-1	21-7-5	24-13-14
Collectors	46-82-72	17-4-8	8-4-4	11-0-2	5-4-6	13-6-8
Small Urban						
Interstate	94-23-0	6-27-12	0-32-41	0-13-24	0-0-13	0-5-10
OFY	--33-82	--33-13	--16-2	--7-0	--3-0	--8-3
OPA	40-33-25	30-18-28	15-8-19	11-6-2	4-8-5	0-27-21
Minor Arterials	67-68-53	3-4-16	5-6-7	14-13-4	7-6-4	4-3-16
Collectors	79-75-75	15-11-13	6-3-5	0-5-4	0-0-0	0-6-3

* The Proportions for each of the three surveyed States are shown in each cell.

Table 7

K-Factor: Summary of Estimated Proportions for Selected Data Element Values *

Area Type and Functional System	≤ 4	5-9	10-14	15-19	19
Rural					
Interstate	0-0-0	1-1-7	99-96-55	0-2-38	0-0-0
OPA	0-0-0	7-1-6	83-58-90	6-36-4	4-5-0
Arterials	0-0-0	5-0-1	85-26-82	7-28-17	3-46-0
Major Collectors	0-0-0	1-1-1	93-29-78	6-64-21	0-6-0
Minor Collectors	0-0-0	0-1-2	100-19-89	0-75-9	0-5-0
Urbanized					
Interstate	0-0-0	21-17-0	79-82-100	0-1-0	0-0-0
OFY	0-0-0	100-24-28	0-72-72	0-4-0	0-0-0
OPA	0-0-0	89-21-48	11-59-52	0-16-0	0-4-0
Minor Arterials	0-0-0	30-6-15	70-59-85	0-32-0	0-3-0
Collectors	0-0-0	24-4-8	76-33-92	0-59-0	0-4-0
Small Urban					
Interstate	0-0-0	0-0-0	100-92-88	0-8-12	0-0-0
OFY	--0-0	--0-19	--100-68	--0-13	--0-0
OPA	0-0-0	96-9-13	4-78-84	0-13-3	0-0-0
Minor Arterials	0-0-0	26-4-12	74-73-88	0-23-0	0-0-0
Collectors	0-0-0	0-7-13	100-33-87	0-60-0	0-0-0

* The proportions for each of the three surveyed States are shown in each cell.

Table 8

Pavement Condition: Proportions for Selected Pavement Condition Ratings *

Area Type and Functional System	0	.1-.9	1.0-1.9	2.0-2.9	3.0-3.9	4.0-4.9	5.0
Rural							
Interstate	0-0-0	0-0-0	14-0-0	24-25-0	28-34-15	34-41-82	0-0-3
OPA	0-0-0	0-0-0	74-8-0	15-18-6	4-60-35	7-14-59	0-0-0
Arterials	0-0-0	0-1-0	63-6-0	19-21-44	13-61-43	5-11-13	0-0-0
Major Collectors	28-0-0	0-1-0	41-4-0	19-46-36	5-32-52	7-17-12	0-0-0
Minor Collectors	88-5-0	0-0-0	10-7-10	2-49-43	0-33-39	0-6-8	0-0-0
Urbanized							
Interstate	0-0-0	0-0-0	9-0-0	48-4-0	20-35-37	23-58-41	0-3-22
OFY	0-0-0	0-0-0	0-0-0	0-18-6	0-60-47	100-22-40	0-0-7
OPA	2-0-0	0-0-3	23-0-2	23-27-36	37-64-50	15-9-9	0-0-0
Minor Arterials	4-0-0	0-0-0	46-5-7	22-44-33	13-30-58	15-21-2	0-0-0
Collectors	12-0-4	0-0-0	17-4-12	13-25-49	30-48-28	28-23-5	0-0-2
Small Urban							
Interstate	0-0-0	0-0-0	0-7-0	4-15-0	56-46-16	40-32-84	0-0-0
OFY	--0-0	--0-0	--0-0	--13-1	--72-16	--15-67	--0-16
OPA	0-0-0	0-0-0	14-0-0	31-25-12	25-63-56	30-11-32	0-1-0
Minor Arterials	9-2-0	0-0-1	8-3-0	13-51-26	60-28-65	10-16-8	0-0-0
Collectors	29-3-0	0-0-0	0-6-7	41-46-52	15-39-37	15-5-4	0-1-0

* The proportions for each of the three surveyed States are shown in each cell.

Table 9

Right Shoulder Width: Summary of Estimated Proportions for Selected Data Element Values *

Area Type and Functional System	No Shoulder	1-4 ft.	5-8 ft.	8 ft.
Rural				
Interstate	0-0-0	0-0-0	7-5-3	93-95-97
OPA	1-3-0	12-19-18	83-62-41	4-16-41
Arterials	7-7-3	59-36-31	33-49-57	1-8-9
Major Collectors	51-12-2	27-53-37	20-29-53	2-6-8
Minor Collectors	91-28-3	8-47-68	1-23-28	0-2-1
Urbanized				
Interstate	0-3-1	0-0-0	2-4-1	98-93-98
OFY	0-4-4	0- -1	1-64-16	99-28-79
OPA	40-16-53	33-10-18	22-57-16	5-17-13
Minor Arterials	32-30-29	40-16-40	25-49-22	3-5-9
Collectors	66-30-24	9-12-50	19-55-24	6-3-2
Small Urban				
Interstate	0-0-0	0-6-0	0-3-0	100-91-100
OFY	--0-0	--0-0	--84-15	--16-85
OPA	42-8-25	11-9-20	38-63-38	9-20-17
Minor Arterials	63-38-39	15-13-29	21-34-29	1-15-3
Collectors	68-32-33	7-23-41	25-28-20	0-17-6

* The proportions for each of the three surveyed States are shown in each cell.

Table 10

Lane Width: Summary of Estimated Proportions for Selected Data Element Values *

Area Type and Functional System	10 ft.	10 ft.	11 ft.	12 ft.	13 ft.	14 ft.	15 ft.	15 ft.
Rural								
Interstate	0-0-0	0-0-0	0-0-0	100-100-100	0-0-0	0-0-0	0-0-0	0-0-0
OPA	0-1-5	0-12-6	1-11-20	99-72-68	0-1-0	0-3-0	0-0-0	0-0-1
Arterials	0-9-43	1-10-19	9-10-29	90-54-8	0-7-1	0-7-0	0-2-0	0-1-0
Major Collectors	17-5-48	8-17-14	22-28-21	50-44-17	1-0-0	0-1-0	2-5-0	0-0-0
Minor Collectors	0-7-71	31-31-14	15-14-7	48-38-2	4-3-0	0-1-5	2-6-0	0-0-1
Urbanized								
Interstate	0-0-0	0-0-0	0-0-0	100-98-100	0-2-0	0-0-0	0-0-0	0-0-0
OFY	0-0-0	0-0-0	0-3-1	100-97-94	0-0-0	0-0-0	0-0-0	0-0-5
OPA	1-0-9	3-3-9	4-0-16	10-14-17	9-1-2	10-1-1	0-1-2	63-80-44
Minor Arterials	0-0-27	12-11-17	2-8-21	44-8-3	7-4-0	7-4-6	0-2-6	28-63-20
Collectors	3-0-46	4-11-17	3-1-8	33-7-6	6-3-2	7-1-0	4-4-1	40-73-20
Small Urban								
Interstate	0-0-0	0-0-0	0-3-0	100-97-100	0-0-0	0-0-0	0-0-0	0-0-0
OFY	--0-0	--0-0	--8-2	--88-98	--4-0	--0-0	--0-0	--0-0
OPA	0-0-11	0-0-22	0-4-15	19-9-37	0-0-0	0-1-1	6-0-0	75-86-14
Minor Arterials	0-0-30	6-6-11	0-4-8	9-14-24	7-4-4	0-5-4	8-3-6	70-64-13
Collectors	0-0-54	0-19-9	0-18-0	35-11-8	9-2-2	3-0-5	25-0-1	28-50-21

*The proportions for each of the three surveyed States are shown in each cell.

When dealing with proportions, the degree of confidence one may have in an estimated proportion is determined by the size of the sample and the error of the estimated proportion. This combination of sample size and sample error limits the minimum proportion to which a desired precision level can be applied. The term "minimum proportion" may be defined as the lowest proportion to which a specified level of confidence in a desired allowable error is valid. The desired allowable error is exceeded for proportions below this critical level. A method for determining this minimum proportion is to take a random subsample of the total estimated proportions and their associated percent relative errors and produce a curve of "best fit" for the relationship between percent relative error and estimated proportion. The curves in Figures 1 and 2 show a curvilinear relationship which is expressed by "best fit" power series curves of the 3rd (Figure 1) and 4th (Figure 2) degrees. Visual inspection shows that the Figure 2 curve is the more representative of the point scatter, and also, that the critical portion of the curve is the portion below 30 percent on the X-axis where there is a rapid rate of increase in the relative error. The Figures 1 and 2 statistic R^2, the coefficient of determination, is the value of the ratio of the variation explained by the curve to the total variation.

The Figures 1 and 2 curves represents a one standard error (68 percent) confidence that any given proportion on the X-axis will not be in error by more than its respective percent value on the Y-axis. As an example, there is 68 percent confidence that the true value of an estimated proportion of 40 percent in Figure 2 will be between 32 and 48 percent ($0.20 \times 40 = 8$; 40 ± 8).

Since the percent relative error in Figures 1 and 2 is expressed in the percent value of one standard error, the minimum proportion for any desired precision level can be calculated easily from the equation:

$$\text{Percent Relative Error in Figure 2} = \frac{\text{Desired Percent Relative Error}}{\text{Z-Value for Desired Confidence Level}}$$

The minimum proportion is the value on the X-axis associated with the value calculated from the above equation for Percent Relative Error, Figure 2.

Z-values for specific confidence levels have been previously listed.

For example, should one wish to know the minimum proportion for a confidence level of 80 percent and an allowable relative error of 20 percent, then

$$\text{Percent Relative Error in Figure 2} = \frac{20}{1.28} = 15.625$$

The corresponding value to 15.625 on the Figure 2 X-axis is about 58 percent, the minimum proportion to which the desired precision level of 80-20 can be applied.

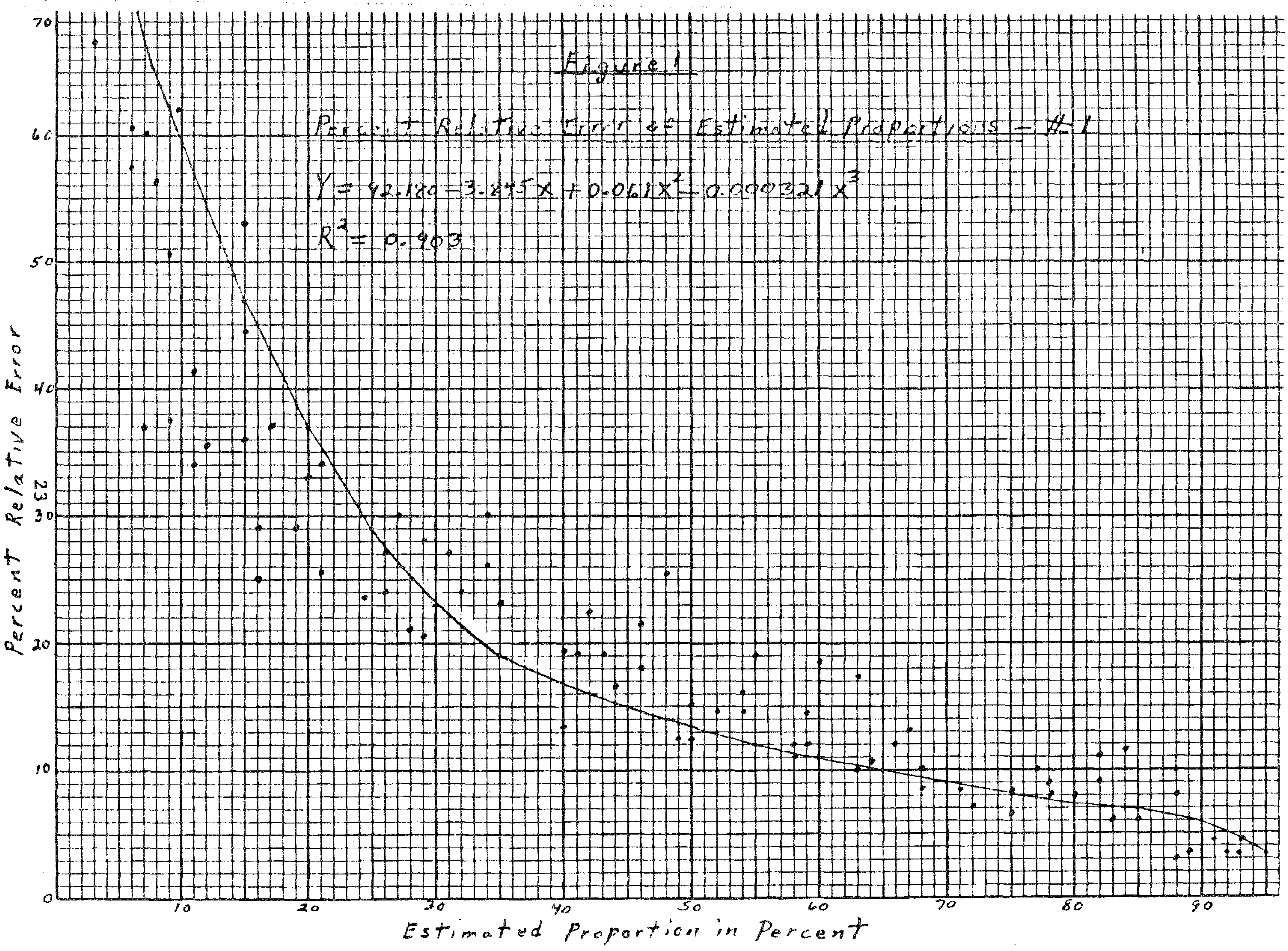


Figure 1

Percent Relative Error of Estimated Proportions — #1

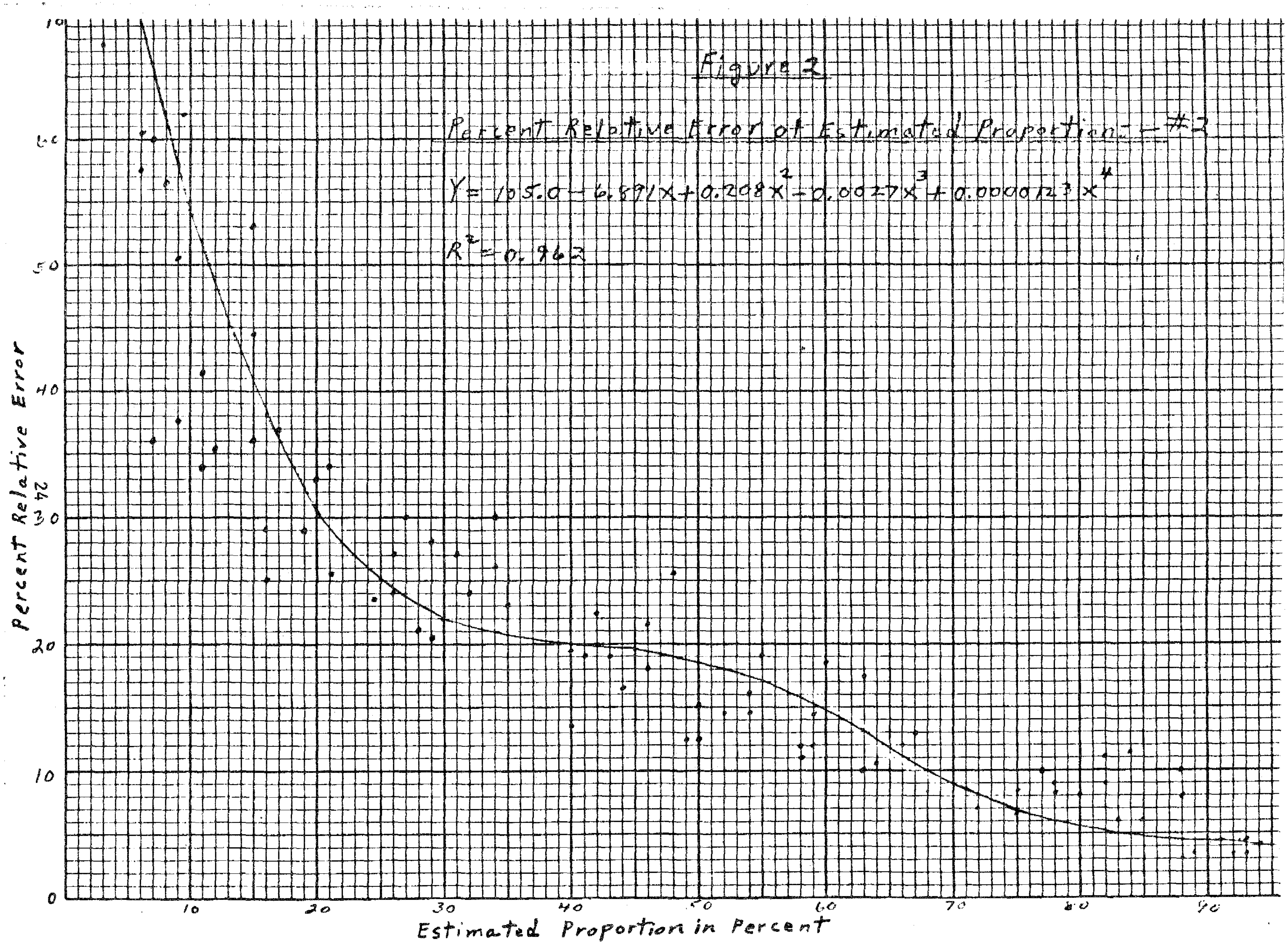


Figure 2
Percent Relative Error of Estimated Proportion — #2
Y = 105.0 − 6.891X + 0.208X² − 0.0027X³ + 0.0000123X⁴
R² = 0.962
Percent Relative Error
Estimated Proportion in Percent
0
10
20
30
40
50
60
70
80
90

Using the method outlined above, the minimum applicable proportions for various precision levels can be calculated for the selected fixed sample panel. Minimum proportions based on the Figure 2 curve are shown on Table 11 for various precision levels:

Table 11

Minimum Proportions Applicable to Selected Precision Levels

Precision Level (Confidence Level-Allowable Percent Error)	Minimum Proportion in Percent
95-10	83
90-10	78
80-10	72
80-15	65
80-20	58
80-25	45
70-25	27
60-25	21
50-25	17

Of interest is the effect of increasing sample size on the minimum proportion applicable to a precision level of 80-10. Doubling the present average sample size of 65 sections per functional system will lower the minimum proportion from the present 72 percent to 56 percent; quadrupling the sample size reduces the minimum proportion to 39 percent. The average sampling rate in this study is about 4.3 percent (65 ÷ 1,500, the average number of road sections sampled per functional system divided by the average number of available road sections per functional system in the three study States).

Appendix D shows the effect of increasing sample size on the value of the minimum proportion for various precision levels. Note that for small proportions a precision level of 80-10 requires sample sizes exceeding 1,000.

The finiteness of the number of sections available for sampling in the various States has the effect of lowering the minimum proportions applicable to desired precision levels. A finite correction factor, $(N - n)/N$, where N is the number of sections available for sampling in a system and n is the number of sections sampled, is effective in fractionally reducing the relative error of an estimated proportion. However, when the estimated proportions are small, 10 percent or less, the sampling ratio n/N must be increased considerably to obtain high precision levels of accuracy. For instance, in a system with 1,500 road sections, it will require 750 sample sections--a 50 percent sampling rate--to have an 80-10 precision level for a 10 percent proportion; or, for the same precision level for a 10 percent proportion, it will require a sample size of 135 sections for a system with 150 sections available--a 90 percent sampling rate. In the above two examples, if the estimated proportion were 40 instead of 10 percent, the required sampling rates will be 14 percent and 63 percent, respectively.

The above statements regarding sampling rates for proportion at specific precision levels are derived from formulas in Appendix E.

As shown in Table 11 and Appendix D, the degree of confidence that one may have in an estimated proportion in this study decreases directly with the size of the proportion and sample size. Table 12 shows the 80 percent confidence level for the lower and upper limits of selected small proportions for average sample sizes (n) of 25, 65, and 115 road sections per functional system.

Table 12

Lower and Upper Limits for Selected Estimated Proportions with Sample Size (n) for an 80 percent Confidence Level

Estimated Proportion (Percent)	Lower Limit		Upper Limit	Lower Limit		Upper Limit	Lower Limit		Upper Limit
20	11.6	-	32.3	14.4	-	27.1	15.6	-	25.2
15	8.0	-	26.6	10.1	-	21.6	11.2	-	19.8
10	4.5	-	20.7	6.1	-	15.9	6.9	-	14.2
5	1.6	-	14.2	2.5	-	9.8	3.0	-	8.3
2.5	0.6	-	10.6	0.9	-	6.4	1.2	-	5.2

The percent relative error at the 80 percent confidence level for any of the data in Table 12 can be calculated as follows:

$$\text{Percent Relative Error} = \frac{\text{(Upper Limit - Lower Limit/2}}{\text{Size of Estimated Proportion}}$$

As an example, the percent relative error for an estimated proportion of 10 percent with a sample size of 65 is:

$$\frac{(15.9 - 6.1)/2}{10} = \frac{4.9}{10} = 49 \text{ percent for an 80 percent confidence level}$$

The formula for the Table 12 data is given in Appendix F.

The monitoring of the fixed sample panel for significant changes at prescribed time periods leads to the question of how effective is this panel in detecting statistically significant differences in proportions at an 80 percent level of confidence. Because of sampling error, what is the smallest percent change that can be considered significant for a given sample size? A formula for answering this question is given in Appendix G. According to the formula, a sample size range of about 60 to 85 road sections per functional system is needed to detect with 80 percent confidence a change of 10 percent as significant. This generally conforms with the average number of road section samples used in this test. The detection of differences of 5 percent as statistically significant requires samples of 240 to 340 sections--a fourfold increase.

APPENDICES

APPENDIX A

Primary Data Elements

Number of Lanes
Degree of Access Control
Median Type
Median Width
Section Length
Grades (Vertical Alignment)
Horizontal Curves (Alignment)
Pavement Type
Pavement Condition
Skid Resistance
Number of Intersections
Number of Bridges
Number of At-Grade RR Crossings
Prevailing Type of Development
Right-of-Way Width
ADT by Time of Day
Average Daily Traffic
Percent Trucks
Peak Hour Parking
Peak Hour Operation
Speed Limit
Shoulder Type
Percent Passing Sight Distance
Terrain
Lane Width
Shoulder Width
Approach Width
Percent Green Time
Urban Location
"K" Factor
"D" Factor
Drainage Condition
Daily Vehicle Miles Traveled

APPENDIX B

Empirical Method for Computing Sample Size

Predetermined ADT volume group strata are assigned areawide (rural, urbanized, or small urban) to each of the five functional systems in a State. The formulas for calculating the sample size, n, for each volume group stratum for a given precision level of accuracy by simple random sampling are:

$$n_h = \frac{n_o}{1 + n_o/N} \; ; \; n_o = \frac{Z^2(s_1^2 + s_2^2)}{d^2}$$

where,

n_h = The required sample size for a given ADT volume group and for a given precision level, corrected for finiteness.

n_o = The required sample size without finite adjustment.

N = The total number of road sections available for sampling in a given volume group for a specific functional highway system in the State.

Z = The value of the normal variate as applied to a specific confidence level and the total number of road sections in a given ADT volume group. It is obtainable for statistical tables.

d = The allowable range of error from the midpoint value of a given ADT volume group. It is expressed as an absolute value and represents the allowable percentage deviation from the midpoint value of the volume group.

s_1^2 = The spatial variance. This refers to the variation of ADT values among road section locations for a given ADT volume group. The square root of this value, s_1, is the spatial standard deviation. The simplest estimator of the standard deviation and its square, the variance, is based on the range of values contained in a volume group stratum, the difference between the largest and smallest limits of a volume group. Analyses show that the normal distribution of ADT values within defined strata (volume groups) can be approximated. Thus, the spatial variance for a volume group can be estimated by the following formula, based on research by L. H. C. Tippett in Biometrika:

$$s_1^2 = \frac{(\text{Range})^2}{12}$$

$$s_1 = \frac{\text{Range}}{3.464} = \text{approximately } 0.30 \text{ of the Range}$$

then,

s_h^2 = The variance of the given volume group.

Range = The value of the difference between the limits of the volume group.

s_2^2 = The temporal variance. This is the variation of ADT over time at a given road section in a given ADT volume group. The square root of this value, s_2, is the temporal standard deviation. The formula for s_2 is:

$$s_2 = (CV)(\overline{X}_h) \text{ and } s_2^2 = (CV)^2(\overline{X}_h)^2$$

where,

CV = The coefficient of variation, a measure of the relative dispersion of individual road section ADT values over time with reference to the midpoint ADT value for a given volume group. Studies based on traffic counting programs have shown that the size of CV varies inversely with traffic volume.[1/] Appendix C shows a relationship between CV and two-way traffic volumes.

$\overline{X}_h$ = The midpoint ADT value of a given volume group. In the computation of temporal variance, the value of CV in Appendix C is referenced to this midpoint value.

An illustration for the computation of sample size for a functional system follows.

To obtain the sample size needed to estimate the quantitative values of selected data elements in a functional system, e.g., Major Collectors, Rural, at a precision level of 80 percent confidence in an allowable error of 10 percent, the following information is available:

Stratum	Predetermined ADT Volume Group	Total Road Sections in Volume Group (N)	Midpoint Value of Volume Group ($\overline{X}$)	Value of $d^2 = (.10\overline{X})^2$	Range of Volume Group (R)
1	0- 2,499	2,326	1,250	15,625	2,500
2	2,500- 4,999	582	3,750	140,625	2,500
3	5,000- 9,999	317	7,500	562,500	5,000
4	10,000-19,999	107	15,000	2,250,000	10,000
5	20,000-29,999	6	25,000	6,250,000	10,000
		3,338			

1/Source: Guide to Urban Traffic Volume Counting, U.S. Department of Transportation, FHWA, October 1975.

Computation, Columns (1) through (6):

Stratum	(1) $s_1^2 = (0.30R)^2$	(2) CV From Appendix C	(3) $s_2^2\ [(CV)(\bar{X})]^2$	(4) $s_1^2 + s_2^2$
1	562,500	0.27	113,906	676,406
2	562,500	0.18	455,625	1,018,125
3	2,250,000	0.14	1,102,500	3,352,500
4	9,000,000	0.11	2,722,500	11,722,500
5	9,000,000	0.0925	5,347,656	14,347,656

Stratum	(5) $n_o = \frac{Z^2(s_1^2 + s_2^2)}{d^2}$; $Z = 1.29$	(6) $n_h = \frac{n_o}{1 + n_o/N}$
1	72.04	70
2	12.05	12
3	9.92	10
4	8.67	8
5	3.82	* 3
	Total sample for functional system =	103

*It is recommended that no less than 3 road sections be sampled for a volume group.

In the above method, the smallest volume group, in this case 0-2,499 ADT, almost invariably requires the largest number of samples because of the large dispersion of the ADT variables. In the event that this number is too large because of cost-manpower limitations, the use of optimum allocation formulas shown below can be considered, particularly for minor collectors in rural areas and collectors in urbanized and small urban areas. It should be noted, however, that under the optimum allocation approach, the desired precision level is achieved only at the total functional system level; whereas, under the recommended method, the desired precision is obtained not only at a stratum level, where it may be needed, but also is upgraded by summation for the overall system estimates.

Optimum allocation formulas:

$$n_o = \frac{Z^2[\Sigma(N_h/N)(s_h)]^2}{d^2} \ ; \ \Sigma \frac{N_h}{N} = 1$$

$$n = \frac{n_o}{1 + n_o/N} \ ; \ n_h = n \frac{N_h s_h}{\Sigma(N_h s_h)}$$

where,

n_o = The sample size required for a given functional system and precision level, not corrected for finiteness.

n = Same as above, but corrected for finiteness.

n_h = Sample size requirement for a given volume group.

N_h = The total number of road sections available in a given volume group.

N = The total number of road sections available in a given functional system.

$s_h^2 = s_1^2 + s_2^2$ = The composite variance (spatial + temporal) for a given volume group.

d = The allowable range of error from the average ADT of the functional system, expressed as an absolute value. This is obtained by weighting the midpoint ADT value of each volume group by its respective total number of road sections.

s_h = The square root of the composite variance for a given volume group.

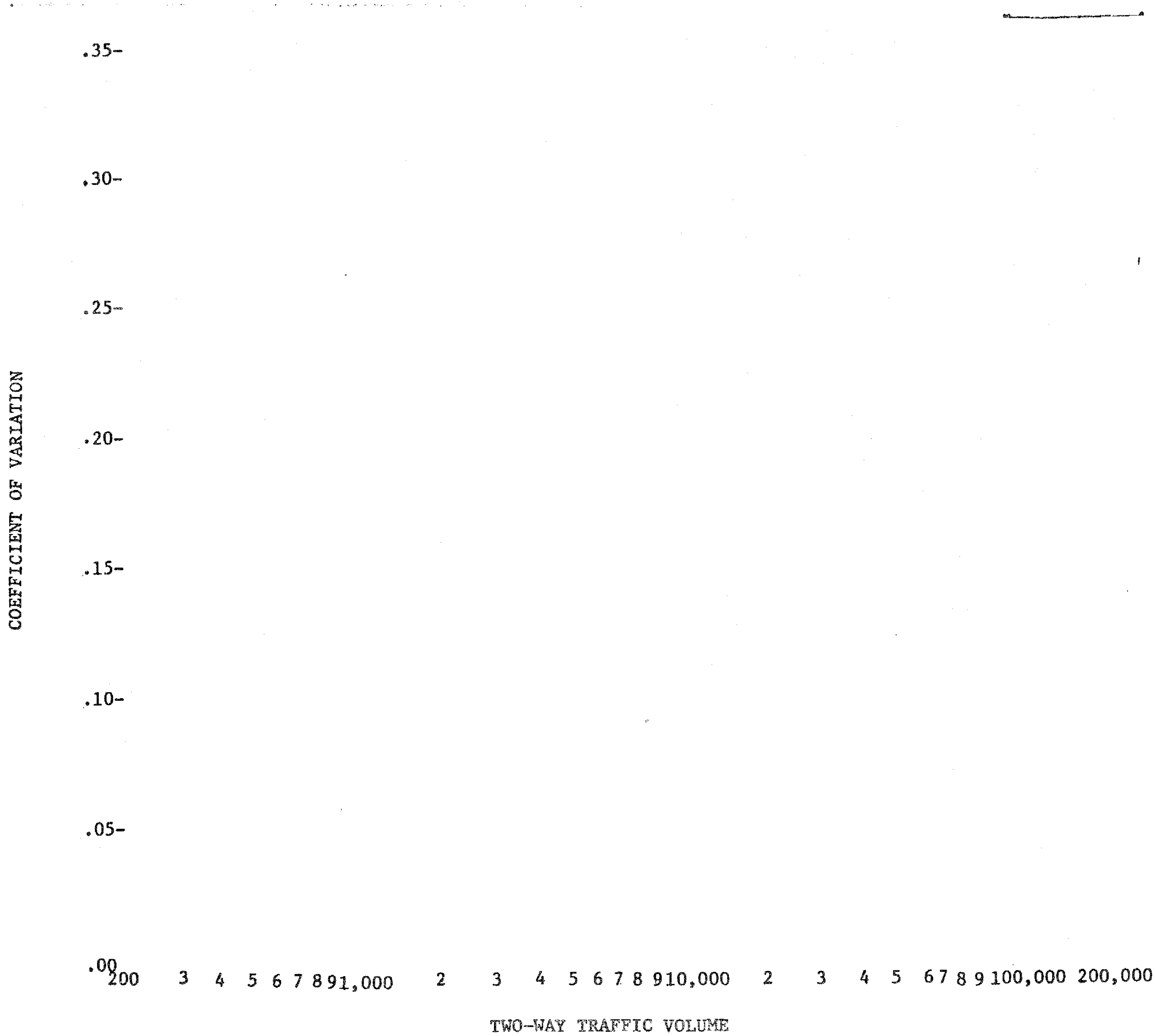
Z = The normal variate = 1.28 for the (80-10) precision level.

Using the optimum allocation formulas for the illustration above, the distribution of sample sizes by stratum for a precision level of 80-10 is:

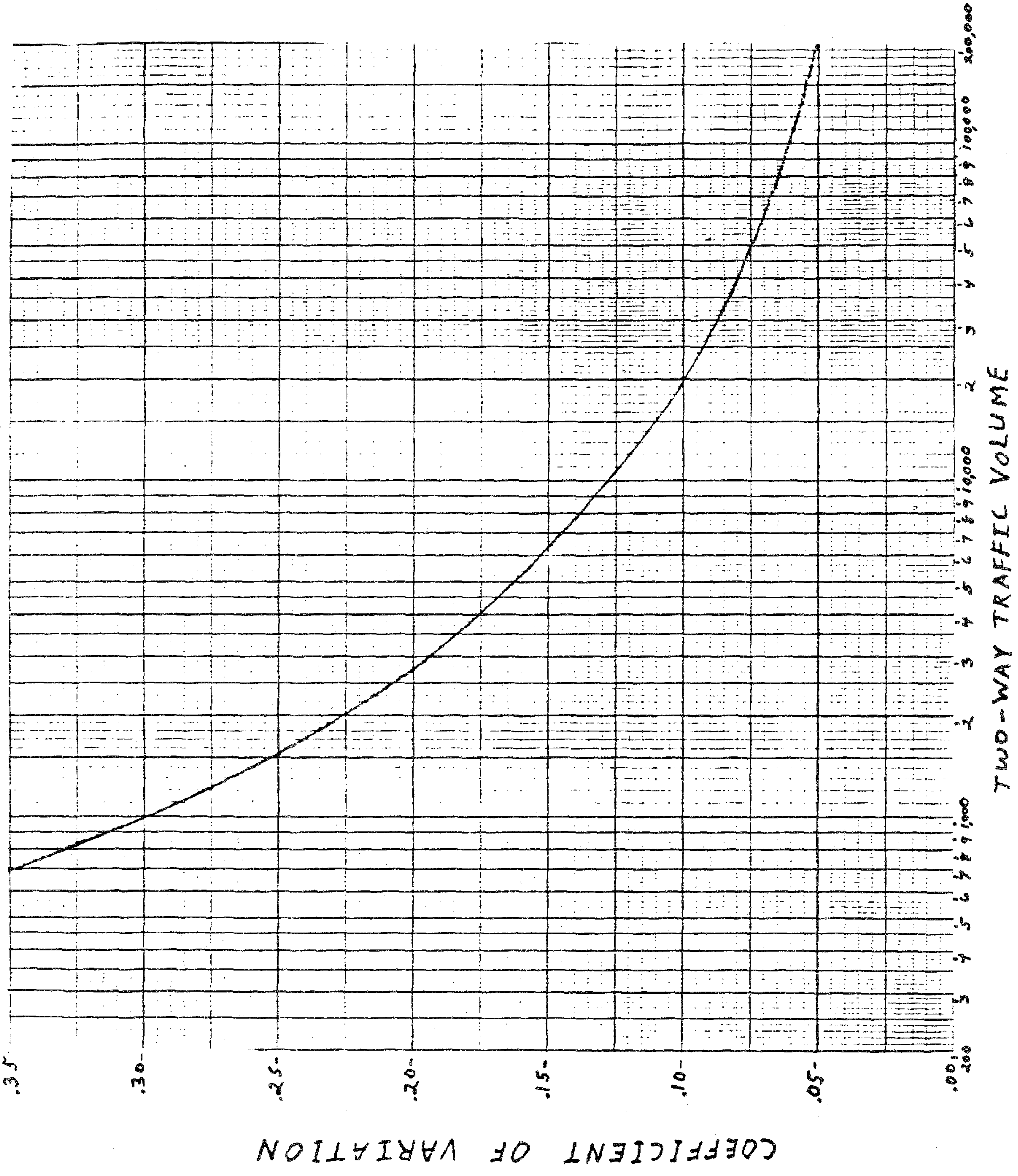
Stratum	Sample Size
1	13
2	4
3	4
4	3
5	3
	27 Functional System Total

APPENDIX C

TEMPORAL TRAFFIC VARIATIONS



APPENDIX C
TEMPORAL TRAFFIC VARIATIONS



APPENDIX D

Value of the Minimum Proportion (In Percent) to Which Desired Confidence Criteria can be Applied - Arranged by Sample Size

Confidence Level %	Relative Error %	Size of Sample						
		30	50	100	250	500	1000	3000
95	10	93	89	80	61	44	28	11
95	15	86	78	64	41	26	15	5
90	10	90	84	73	52	35	21	8
90	15	80	70	55	32	19	11	4
80	10	84	76	62	40	25	14	5
80	15	71	59	42	23	13	7	2
68	5	93	89	80	61	44	29	12
68	7.5	86	78	64	41	26	15	6
68	10	77	67	50	29	17	19	3
68	15	59	47	31	15	8	4	2
68	20	45	33	20	9	5	3	1

The formulas for calculating the minimum proportion (p) values are:

$$q/p = d^2/t^2 \cdot n \quad \text{and} \quad p = \frac{100}{1 + q/p}$$

where,

d = Percent allowable error.

t = Normal variate for the desired confidence level.

p = Minimum proportion in percent.

q = 1 - p

n = Number of sampled sections.

APPENDIX E

Formula for Required Sampling Rates for Specific Proportions

Given, the formula:

$$d^2 = (N - n)/N \cdot \frac{t^2 \, q/p}{n}$$

and, solving for n

$$n = \frac{t^2 \, q/p \cdot N}{d^2 N + t^2 \, q/p}$$

then,

the sampling rate = n/N

where,

n = Number of sections to be sampled.

d = Percent allowable error.

t = The normal variate for the desired confidence level.

N = Total number of sections available for sampling.

p = The estimated proportion under consideration.

q = 1 - p

APPENDIX F

Lower and Upper Limits of Estimated Proportions

The formula for the lower and upper limits of the estimated proportions presented in Table 12 is derived from a modification of the general quadratic formula:

$$\text{Upper or Lower Limit of } p = \frac{b \pm (b^2 - 4ac)^{1/2}}{2a}$$

where,

p = The true proportion.

$\hat{p}$ = The estimated proportion.

$a = (n^2 + t^2 n)$

$b = (2nx + t^2 n)$; $x = \hat{n}p$

$c = (\hat{n}p)^2 = x^2$

n = Sample size.

t = Value of the normal variate for the desired confidence level.

APPENDIX G

Formulas for Sample Size Required (n) to Detect A Significant Difference Between Estimated Proportions

Given the formula:

$$(p_2 - p_1)^2 = Z^2[\bar{p}\bar{q}(\frac{1}{n_1} + \frac{1}{n_2})]$$

Assuming that the sample size n_1 and n_2 are the same in a fixed sample panel, then $n_1 = n_2$; let $\bar{p} = p_1 + p_2/2$ and $2/n_o = \frac{1}{n_1} + \frac{1}{n_2}$.

Solving for n_o,

$$(p_2 - p_1)^2 = Z^2(\bar{p}\bar{q} \cdot 2/n_o)$$

$$n_o = \frac{2(Z)^2(\bar{p}\bar{q})}{(p_2 - p_1)^2} \quad \text{and } n = \frac{n_o}{1 + n_o/N}$$

where,

n = The required sample size, corrected for finite population, needed to detect a significant difference between the estimated proportions p_1 and p_2 at a given level of confidence.

n_o = Same as above, but not corrected for finite population.

n_1 and n_2 = The number of road sections in the sample panel for time periods #1 and #2.

p_1 = The estimated proportion for a given data element attribute on a functional system at time period #1.

p_2 = The same as above for time period #2.

$\bar{p}$ = $(p_1 + p_2)/2$

$\bar{q}$ = $1 - \bar{p}$

Z = The normal variate for a given level of confidence.

N = The total number of road sections available for sampling in a functional system.

Example:

What sample size (n) is required to be 80 percent certain of detecting a statistically significant difference between two estimated proportion $p_1 = 0.65$ and $p_2 = 0.55$ on a functional system having 200 road sections?

Let,

$$\bar{p} = (0.55 + 0.65)/2 = 0.60$$

$$\bar{q} = 1 - 0.60 = 0.40$$

$$Z = 1.29$$

$$N = 200$$

then,

$$n_o = \frac{2(1.29)^2\ (0.60)(0.40)}{(0.55 - 0.65)^2} = \frac{0.799}{0.01} = 80$$

$$n = \frac{80}{1 + 80/200} = 57$$