VOL. 30, NO. 11 DECEMBER 1959

Public Roads

A JOURNAL OF HIGHWAY RESEARCH

PUBLISHED
BIMONTHLY BY THE
BUREAU OF
PUBLIC ROADS,
U.S. DEPARTMENT
OF COMMERCE,
WASHINGTON

The Pacific Highway (Interstate Route 5) south of Salem, Oregon.

Public Roads

A JOURNAL OF HIGHWAY RESEARCH

Published Bimonthly

Vol. 30, No. 11 *December 1959*

C. M. Billingsley, *Editor*

IN THIS ISSUE

U.S. DEPARTMENT OF COMMERCE

FREDERICK H. MUELLER, Secretary

BUREAU OF PUBLIC ROADS

BERTRAM D. TALLAMY, Administrator

ELLIS L. ARMSTRONG, Commissioner

BUREAU OF PUBLIC ROADS

Washington 25, D.C.

REGIONAL OFFICES

No. 1. Delaware and Hudson Bldg., Albany 1, N.Y.
Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Rhode Island, and Vermont.

No. 2. 707 Earles Bldg., Hagerstown, Md.
Delaware, District of Columbia, Maryland, Ohio, Pennsylvania, Virginia, and West Virginia.

No. 3. 50 Seventh St. NE., Atlanta 23, Ga.
Alabama, Florida, Georgia, Mississippi, North Carolina, South Carolina, Tennessee, and Puerto Rico.

No. 4. South Chicago Post Office, Chicago 17, Ill.
Illinois, Indiana, Kentucky, Michigan, and Wisconsin.

No. 5. 4900 Oak St., Kansas City 12, Mo.
Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, and South Dakota.

No. 6. Post Office Box 12037, Ridglea Station, Fort Worth 16, Tex.
Arkansas, Louisiana, Oklahoma, and Texas.

No. 7. New Mint Bldg., San Francisco 3, Calif.
Arizona, California, Nevada, and Hawaii.

No. 8. 740 Morgan Bldg., Portland 8, Oreg.
Idaho, Montana, Oregon, and Washington.

No. 9. Denver Federal Center, Bldg. 40, Denver 2, Colo.
Colorado, New Mexico, Utah, and Wyoming.

No. 10. Post Office Box 1961, Juneau, Alaska.
Alaska.

No. 15. 1440 Columbia Pike, Arlington, Va.
Eastern National Forests and Parks.

PUBLIC ROADS is sold by the Superintendent of Documents, Government Printing Office, Washington 25, D.C., at $1 per year (50 cents additional for foreign mailing) or 20 cents per single copy. Subscriptions are available for 1-, 2-, or 3-year periods. Free distribution is limited to public officials actually engaged in planning or constructing highways, and to instructors of highway engineering. There are no vacancies in the free list at present.

Use of funds for printing this publication has been approved by the Director of the Bureau of the Budget, March 28, 1958.

Circular Reinforced Concrete Columns Subjected to Direct Stress and Bending

Prepared for the Bureau of Public Roads
by A. ANTHONY TOPRAC [1]

Based on the American Association of State Highway Officials' bridge specifications for allowable loads, a procedure for designing eccentrically loaded round columns has been developed by the author. Under the enlarged highway construction program, bridge engineers will more frequently be confronted with the problem of designing round columns for bridge pier bents.

The principal benefit of this article will be derived from the method outlined in the text, which will assist bridge designers who may want to expand the graphs presented here or make similar graphs for different bridge column sections. Plotted curves are given for a range of 24- to 60-inch diameter columns in 6-inch intervals and for three quantities of reinforcing steel. In design problems for which the graphs are applicable, considerable tedious work will be avoided.

THE 1957 American Association of State Highway Officials' specifications (*1*) [2] on the design of eccentrically loaded columns state that for ratios of eccentricity to depth (diameter) not greater than 0.5, the allowable load could be calculated on the basis of uncracked section, and for ratios greater than 0.5 the design or the allowable load should be based on the theory of cracked sections.

Based on these specifications, a procedure for designing eccentrically loaded columns has been developed and is presented here. Curves have also been prepared for 24- to 60-inch diameter columns in 6-inch intervals for three percentages of steel in each case.

The problem of designing or investigating round columns subject to bending is more difficult than the same problem for rectangular columns. With the increased activity in highway construction, many engineers will be confronted with the design of round columns for bridge pier bents. To facilitate the design and investigation of such columns, a method is presented in the following discussion. The method hinges on the ability of the designer to guess at the position of the neutral axis; the designer then calculates the allowable stress and makes proper use of the expression $N/A_t \pm Mc/I_t$. Once the position of the neutral axis is checked, properties of the cross section such as r^2 and K are calculated and used to obtain allowable stresses.[3]

Approaches to Problem of Eccentrically Loaded Columns

The author found three methods which are used to solve the problem of eccentrically loaded columns. Although these methods are based on the same fundamental laws of equilibrium, on the surface they look different. Consequently, it is worthwhile to outline them in short form.

1. In the first of these methods, used by Boguslavsky (*2*), the effective portions of a cracked round section are a circular segment of concrete in compression and a circle of steel bars; some of them in compression, others in tension. The internal stresses acting on the transformed section vary elastically from f_s/n to f_c in straight-line proportions. The resultant C of the compressive stresses acting on the circular segment of concrete is represented by the volume of the ungula (wedge cut by a sloping plane from a circular prism), while the tensile resultant T of the tensile and compressive stresses acting on the concrete ring, which replaces the steel, is represented by the volume obtained by multiplying the area of the ring by the surface of the circular prism, intercepted between the horizontal and sloping planes. Formulas are used for C and T, for their moment arms from the axis of the column, and for the moments. The derivation of these formulas, which contain one-half the angle subtended at the center of the circle by the chord on the neutral axis, may be found in Peabody's *Design of Reinforced Concrete Structures* (*3*).

2. In the second method, presented by Dunham (*4*), an analysis of the stress diagram shows that the imaginary solid that represents the compression upon the concrete is an ungula of a circular cylinder (a wedge-shaped solid). In order to simplify the calculations, the longitudinal rods in the column are assumed to be replaced by a hollow steel shell or pipe which has the same cross-sectional area as that of all of the longitudinal reinforcement. This steel shell is subjected to compression upon part of its area and to tension upon the rest of its cross section. Therefore, the diagrams that represent the stresses upon these two parts of the ring will be like hollow ungulas or like wedges cut from a pipe. In order to have equilibrium, it is obvious that $\Sigma V=0$ and $\Sigma M=0$. Curves prepared by Dunham give the volume and the moment of ungula about the neutral axis for the concrete in compression. Similar curves give the volume and moment of wedge of hollow cylinder (for the steel) about the same neutral axis. Then, the desired allowable stresses are obtained by applying the two equations of equilibrium.

3. The third method, presented by Large (*5*), involves the following computations:

(1) Calculate trial stresses by $N/A_t \pm Mc/I_t$, assuming no area lost by cracking, and from the stress diagram determine the approximate depth of the crack.

(2) Compute corrected stresses by $N/A_t \pm Mc/I_t$, of the cracked section, using curves developed by Professor J. R. Shank for the properties of circular segments. The curves (*6*) showing these properties are presented in figure 1.

Approach Adopted

Of the three methods the last one was found to be relatively more practicable and was adopted in the preparation of curves plotting load versus eccentricity. To facilitate computations, figure 2 was prepared.

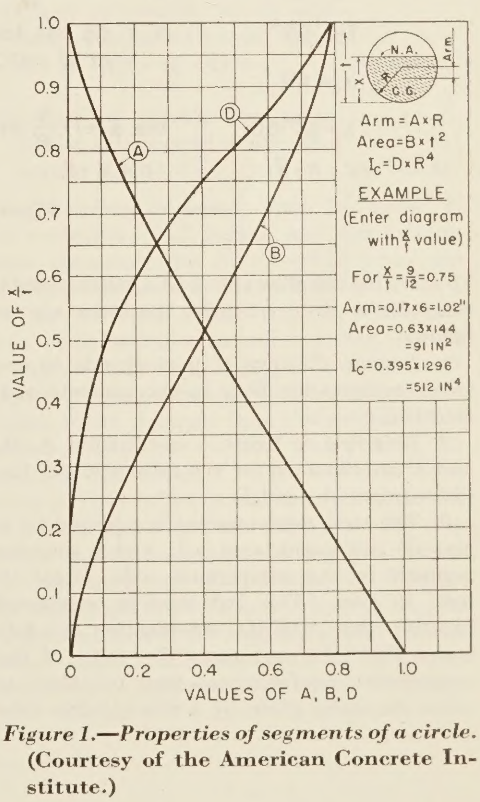


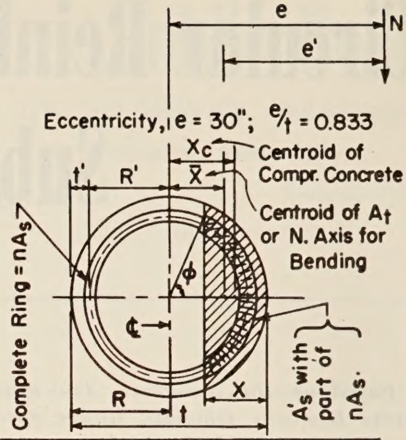
***Figure 1.—Properties of segments of a circle.* (Courtesy of the American Concrete Institute.)**

[1] This article was prepared by Professor Toprac, Associate Professor of Civil Engineering, University of Texas, while employed during the summer months of 1958 by the Bureau of Public Roads, Texas Division office, Austin.

[2] Italic numbers in parentheses refer to the list of references on p. 245.

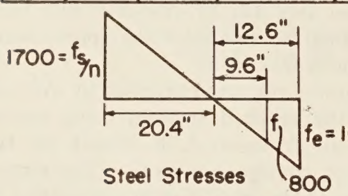
[3] Letter symbols used in this article are defined on pp. 244–245.

Given: $t = 36''$ $t' = 3''$ $d = 33''$ $A_s = 12.7 \text{ in.}^2$ $A_g = 1018 \text{ in.}^2$ $p = 0.0125$ $n = 10$ Eccentricity, $e = 30''$; $e/t = 0.833$



N.A. from Comp. Edge	$k = \frac{X}{t}$	Eff. Area of Concrete		Area of Steel			Transf. Area A_t	$\frac{N}{A_t}$	Coeff. A	X_c	Bending N.A. $\bar{X}$	$X_c - \bar{X}$
		Coeff. B	A_c	nA_s	$\lambda . A_s = A'_s$	A_{st}						
X = 10.8	0.30	0.20	259	127	0.34×127 = 43	170	429	233	0.65	11.72	8.30	3.42
= 11.5	0.32	0.218	282	127	0.354×127 = 45	172	454	220	0.624	11.25	8.20	3.05
= 12.6	0.35	0.25	324	127	0.38×127 = 48	175	499	200	0.59	10.64	8.00	2.64

Moment of Inertias							Concr. c	$\frac{Ne'c}{I_t}$	$\frac{Ne'(R'+\bar{X})}{I_t}$	Combined Stresses		Calc. X = $d \times \frac{f_c}{f_c + f_s/n}$	$r^2 = \frac{I_t}{A_t}$	$K = \frac{t^2}{2r^2}$	f_a	f_e
Coeff. D	I_c	I_s	$A_c(X - \bar{X})^2$	$nA_s\bar{X}^2$	$A'_s(X' - \bar{X})^2$	$\Sigma = I_t$				f_c	f_s/n					
0.02	2,100	14,200	3,000	8,750	690	28,740	9.7	733	1760	966	1527	12.8				
0.025	2,530	14,200	2,600	8,550	720	28,600	9.8	746	1770	966	1550	12.7				
0.034	3,580	14,200	2,250	8,125	660	28,815	10.0	763	1755	963	1555	12.6	57.7	11.2	628	1105



Steel Stresses

$f_s = 17{,}000$

$f'_s = 2nf = 16{,}000$

$$N = 100{,}000 \times \frac{f_e}{f_c} = 100{,}000 \times \frac{1050}{963}$$

$= 109{,}000$ lbs.

For $X = 10.8''$

$\cos \phi = 1 - \frac{7.8}{15} = 0.48$

$\phi = 61.3°$

$259 \times 11.72 = 3030$

$43 \times 12.3 = \frac{530}{3560}$

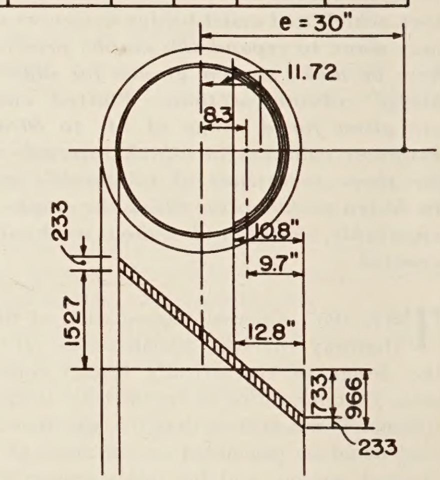
$\bar{X} = \frac{3560}{429} = 8.3''$

$X_c - \bar{X} = 11.72 - 8.30 = 3.42$

$X' - \bar{X} = 12.3 - 8.3 = 4.0$

$C = R - \bar{X} = 18 - 8.3 = 9.7''$

$e' = 30 - 8.3 = 21.7''$



Formulas

$A_c = Bt^2$ $\bar{X} = \frac{\Sigma(\text{Areas}) \times (\text{dist.})}{A_t}$

$I_c = DR^4$ $X' =$ c.g. dist. from ¢ of $A' = 57.3 R' \frac{\sin\phi}{\phi°}$

$X_c = AR$

$\lambda = \phi°/180$ $\cos.\phi = 1 - \frac{a}{R'}$ where

$c = R - \bar{X}$ $a = X - t'$

Figure 2.—Sample calculations for design graphs (figs. 3–6).

To solve for the allowable load a column could carry for a given e/t ratio, the steps are as follows:

1. Assume distance of neutral axis to extreme compression fiber in the concrete and calculate $k = x/t$.

2. Referring to figure 1, coefficients A, B, and D are obtained for the calculation of the distances x_c, A_c, and I_c.

3. The steel reinforcement is replaced by a ring of equivalent area nA_s and a circular segment in the compression side, equal to nA'_s in area. This last steel is considered because the AASHO specification 1.7.8(f) states that "In such cases the value of the compressive reinforcement may be taken as twice the value given by a straight-line relationship between stress and strain and the modular ratio, n . . ."

4. Find the direct stress N/A_t, assuming $N = 100{,}000$ pounds.

5. Find $\bar{x}$ for the neutral axis for bending only, by finding the center of gravity of the transformed area and calculating the moment of inertia of the transformed area to be used in the expression Mc/I.

6. Calculate the actual stresses for 100,000 pounds eccentric load,

$$f_c = \frac{Mc}{I_t} + \frac{N}{A_t} \text{ and } f_s/n = \frac{-M(R' + \bar{x})}{I_t} + \frac{N}{A_t}.$$

7. Calculate the allowable compressive stress f_e and check the stresses in the tension and compression steel that correspond to this calculated f_e.

8. The eccentric allowable load then is $N = 100{,}000 \times$ (allowable compressive stress/actual compressive stress).

The above procedure was used in the preparation of the curves given in figures 3–6.

Practical Illustrations

The following are illustrative examples which show the various steps just mentioned.

Example 1.—Given a 36-inch outside diameter tied column with 10 No. 10 bars as longitudinal reinforcement, calculate the allowable compressive load applied 5 inches from the center of the column, if $f'_c = 3{,}000$ p.s.i. and $n = 10$.

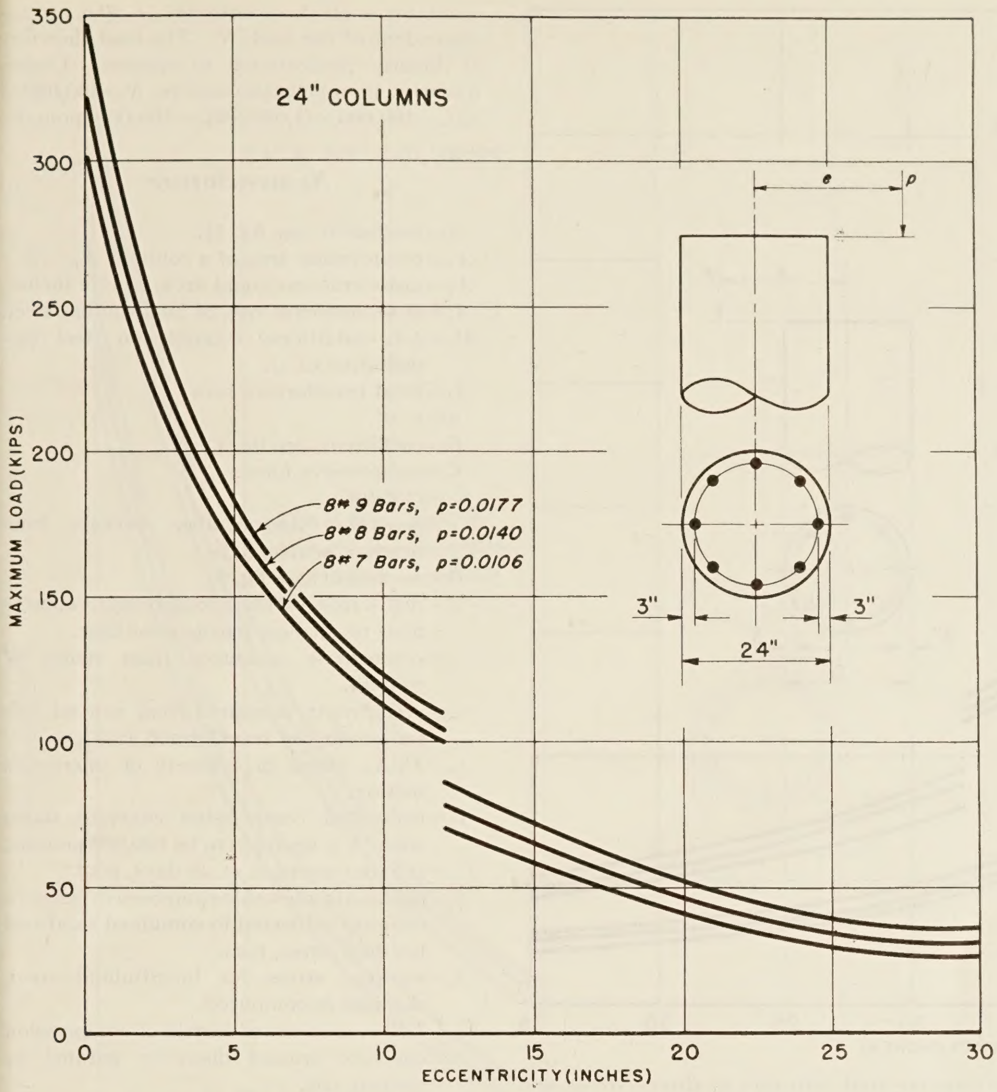


Figure 3.—Maximum load for 24-inch circular-tied columns in direct stress and bending.

Solution.—Since $e/t=5/36=0.139$ which is less than 0.5, use of the uncracked section is allowed by the specifications.

$P_t=0.8\ (0.225\ f'_cA_g+A_sf_s)=0.8\ (0.225\times 3{,}000\times 1{,}018+12.7\times 16{,}000)=0.8\ (686{,}000+203{,}000)=713{,}000$ pounds centrally applied load.

The transformed area, $A_t=1.018+9\times 12.7=1{,}133$ in.^{2}

The concrete compressive stress, $f_a=713{,}000/1{,}133=628$ p.s.i.

Then $C_1=f_a/0.40f'_c=628/1{,}200=0.523$.

The moment of inertia of the transformed area, $I_t=I_c+I_s$. Where $I_c=\text{area}\times t^2/16=1{,}018\times 1{,}296/16=82{,}500$ in.^{4}, and $I_s=\text{area}\times (t-2t')^2/8=115\times 30^2/8=13{,}000$ in.^{4} Thus, $I_t=82{,}500+13{,}000=95{,}500$ in.^{4}

$r^2=I_t/A_t=95{,}500/1{,}133=84.0$, and $K=t^2/2r^2=1{,}296/2\times 84=7.70$.

$C_1K=0.523\times 7.70=4.03$, and $1+C_1K(e/t)=1+4.03\times 0.139=1.560$.

$N=P_t/1+C_1K(e/t)=713{,}000/1.560=457{,}000$ pounds.

Example II.—Same as example I, except $e=30$ inches.

Solution.—Since $e/t=30/36=0.833$ which is greater than 0.5, use method of cracked section. To facilitate calculations, figure 2 is used. Assume $kt=10.8$ inches and $k=0.30$. From figure 1 coefficients B, A, and D are found to be 0.20, 0.65, and 0.02, respectively. The compression concrete area then is $0.20\times 36^2=259$ in.^{2} The equivalent concrete that replaces the reinforcing bars is a ring of $n\times A_s=10\times 12.7=127$ in.^{2}

Since the AASHO specification 1.7.8(f) states that ". . . the value of the compressive reinforcement may be taken as twice the value given by a straight-line relationship between stress and strain and the modular ratio . . .", a circular segment of the equivalent concrete ring is added to the remaining area. A factor, λ, is calculated for this purpose based on the angle $\phi°$ (half of the angle subtended at the center of the column by chord which coincides with the assumed neutral axis). This angle is given by the expression $\cos\phi=1-a/R'=1-7.8/15=1-0.52=0.48$; then the angle $\phi=61.3°$. The amount of additional compressive steel is $\phi°/180\times nA_s=0.34\times 127=43$ in.^{2}, and all the transformed area of the cracked section becomes equal to 429 in.^{2} Assuming the load on the column to be $N=100{,}000$ pounds, the direct compression stress is $N/A_t=233$ p.s.i.

The same transformed section is also in bending because of this load N, which is applied eccentrically. Bending takes place around a neutral axis (for bending only) which passes through the center of gravity of $A_t=429$ in.^{2}

The center of gravity of the concrete area is 0.65×18 inches$=11.72$ inches from the center of the column. By taking the first moments of the areas, the bending neutral axis is found to be $\bar{x}=8.30$ inches away from the center of the column. The moment of the concrete is 3,030 in.^{3} and that of the additional compressive steel is 530 in.^{3} The moment arm of this steel is obtained by the expression, $x'=(57.3R')\sin\phi/\phi°=(57.3/61.3°)(15)(\sin 61.3°)=14\times 0.877=12.3$ inches.

To calculate bending stresses, the formula Mc/I is used where the moment of inertia $I=I_t$. The moment of inertia for the transformed area around the bending neutral axis, I_t, is calculated as follows:

The moment of inertia, I, of compressive concrete around its center of gravity is, $I_c=0.02\times 18^4=$	2, 100
Additional moment of inertia, I, to transfer above to neutral axis equals $259\times 3.42^2=$	3, 000
Moment of inertia, I, of the equivalent ring, $I_s=nA_s\times(2R')^2/8=127\times 30^2/8=$	14, 200
Additional moment of inertia, I, to transfer above to neutral axis equals $nA_s(\bar{x})^2=127\times 8.3^2=$	8, 750
Moment of inertia, I, of additional compressive steel equals $A'_s\ (x'-\bar{x})^2=43\times 4.0^2=$	690
Moment of inertia for transformed area around the bending neutral axis, $I_t=$	28, 740 in.^{4}

The bending moment then is equal to $Ne'=100{,}000\times 21.7$ inches, and the extreme fiber concrete stress $Ne'c/I_t=733$ p.s.i.; the bending tensile stress equals 1,760 p.s.i.

The combined stresses are $733+233=966$ p.s.i. compression in the concrete and $-1{,}760+233=-1{,}527$ p.s.i. tension in the steel. However, these stresses require that $kt=12.8$ inches which is larger than the assumed 10.8 inches. Therefore a new value must be assumed and checked. This has been done twice and the values appear in figure 2. It will be noticed that the second guess of $x=kt=11.5$ inches resulted in a calculated $x=12.7$ inches. Having these two trial values, the designer can obtain the exact value of x by plotting the difference between the assumed and calculated values against the assumed x as abscissas. Such a plot indicates that $x=12.6$ inches,

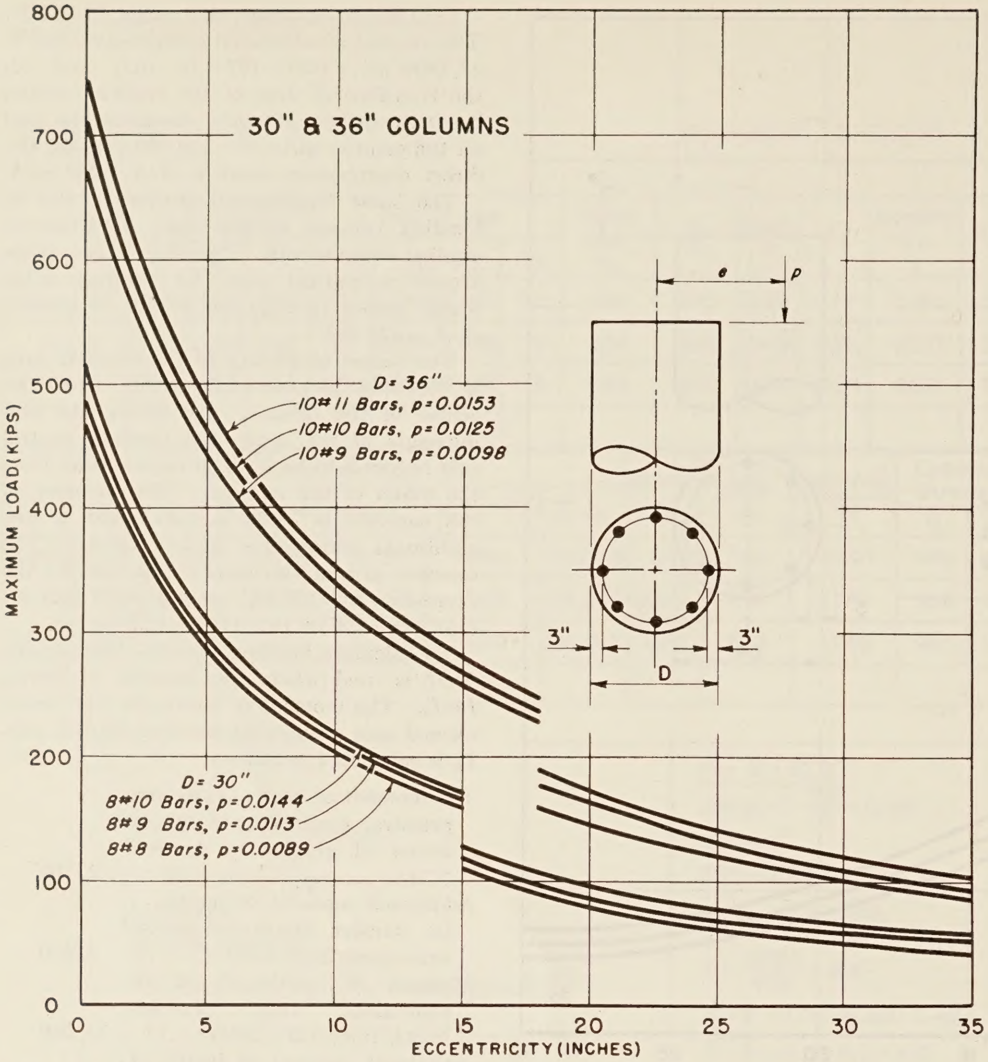


Figure 4.—Maximum load for 30- and 36-inch circular-tied columns in direct stress and bending (f'_c=3,000 p.s.i.; n=10; intermediate, hard and rail steel grade reinforcement).

which when used in the third trial, checks out.

After the value of x has been determined, the C_1 and K quantities and f_a=628 p.s.i. (see example I) are used to obtain the allowable compressive stress. Then,

$$f_e=f_a\left(\frac{1+K(e/t)}{1+C_1K(e/t)}\right)\text{--AASHO spec. 1.7.8(f)}$$

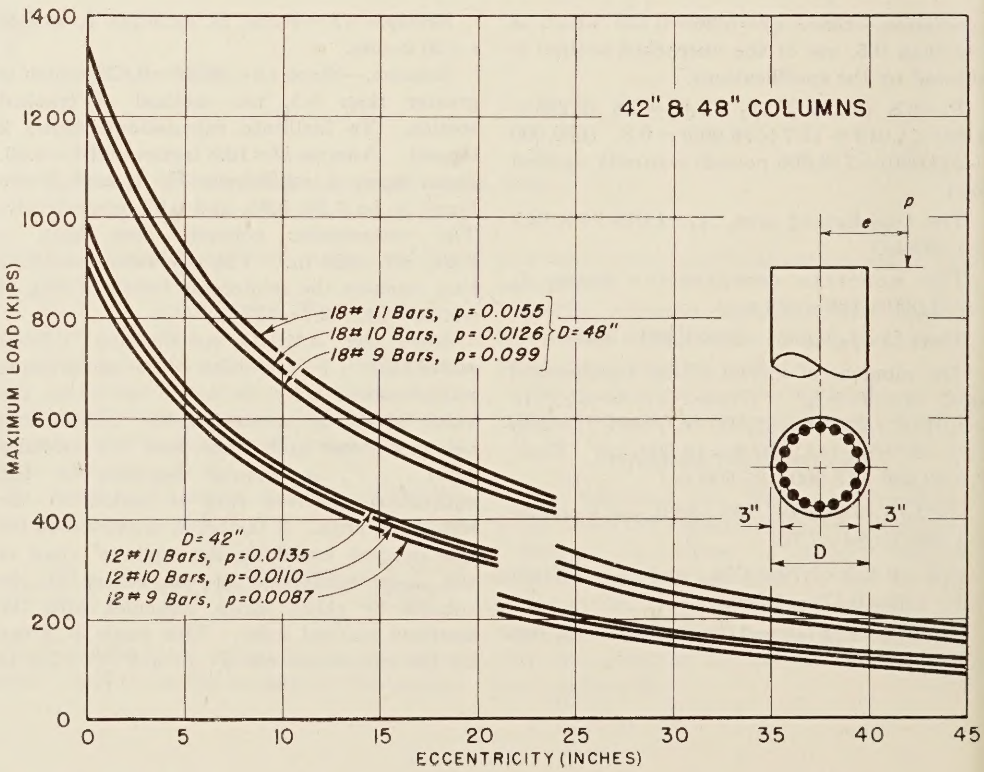
$$=628\left(\frac{1+11.2\times 0.833}{1+0.523\times 11.2\times 0.833}\right)$$

$$=628\times\frac{10.33}{5.88}=1{,}105 \text{ p.s.i.}$$

A check for the stress of the compression steel shows that it is critical; and if f_e=1,105 p.s.i. is used, $f'_s/2n$ will exceed the allowable value of 800 p.s.i. ($f'_s=2n$ × stress in equivalent concrete ≦16,000). If the value of 800 p.s.i. is used, then f_e is reduced to 800× (12.6/9.6)=1,050 p.s.i. The stress in the tension steel becomes 17,000 p.s.i., and in the compression steel, 16,000 p.s.i.

For any column cross section, the properties of the cracked section (I_t, c, $\bar{x}$, etc.) are constant for a given eccentricity, e, and are independent of the load, N. The load therefore is directly proportional to stresses. Consequently, the allowable load is, N=100,000× f_e/f_c=100,000×(1,050/963)=109,000 pounds.

Figure 5 (right).—Maximum load for 42- and 48-inch circular-tied columns in direct stress and bending (f'_c=3,000 p.s.i.; n=10; intermediate, hard and rail steel grade reinforcement).



Nomenclature

A=coefficient (see fig. 1).
A_c=compression area of a column, $A_c=Bt^2$.
A_g=gross cross-sectional area, square inches.
A_s=cross-sectional area of longitudinal steel.
$A'_s=\lambda A_s$=additional compression steel (see definition of λ).
A_t=total transformed area.
$a=x-t'$.
B=coefficient (see fig. 1).
C=compressive force.
$C_1=f_a/0.40f'_c$.
c=concrete extreme fiber distance from bending neutral axis.
D=coefficient (see fig. 1).
d=depth from extreme compression concrete fiber to extreme tensile steel fiber.
e=eccentricity measured from center of column.
e'=eccentricity measured from neutral axis for bending of transformed area.
$f_a=P_t/A_t$, stress in concrete of uncracked section.
f_c=computed compressive concrete stress when N is assumed to be 100,000 pounds.
f'_c=cylinder strength at 28 days, p.s.i.
f_e=maximum allowable compressive stress in columns subjected to combined axial and bending stress, p.s.i.
f_s=working stress for longitudinal steel, nominal or computed.
$I_c=DR^4$= moment of inertia of compression concrete around diameter parallel to neutral axis.
I_s=moment of inertia of steel ring around diameter.

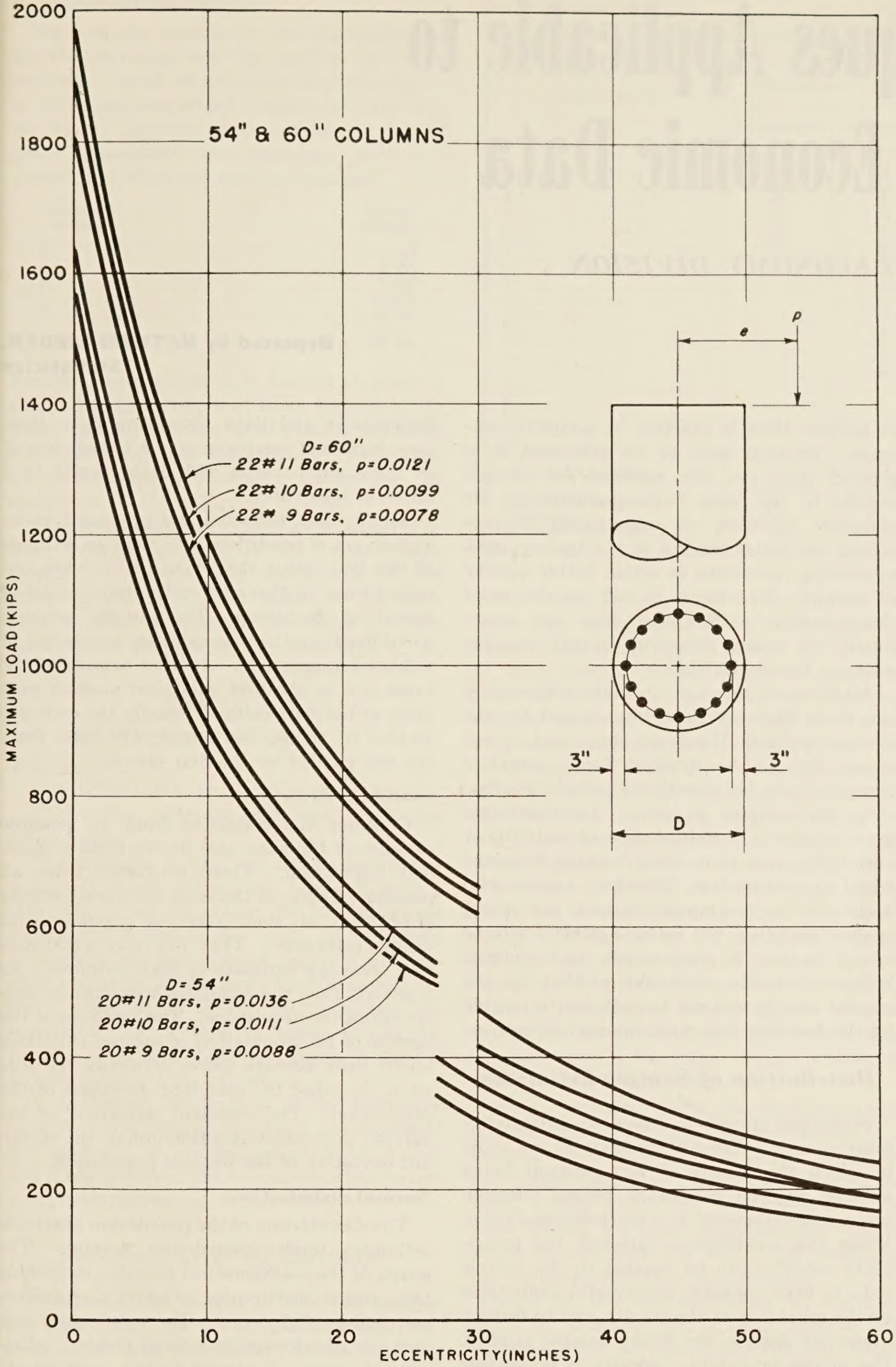


Figure 6.—Maximum load for 54- and 60-inch circular-tied columns in direct stress and bending (f'_c=3,000 p.s.i.; n=10; intermediate, hard and rail steel grade reinforcement).

I_t = total moment of inertia around neutral axis for bending.

$K = t^2/2r^2$, a factor used in the design of eccentrically loaded columns.

$k = x/t$, ratio of depth of compression area to the diameter of column.

M = bending moment.

N = eccentric load.

n = ratio of modulii of steel to concrete.

P_t = total load on tied column centrally applied, pounds.

p = ratio of longitudinal steel area to gross column area.

R = radius of column = $t/2$.

R' = radius of steel ring = $R - t'$.

r = radius of gyration of transformed section.

t = diameter of column.

t' = amount of concrete cover for longitudinal reinforcement.

$x = kt$, depth of compression area.

$x_c = AR$ = distance from center of column to c.g. of compression concrete.

$\bar{x}$ = distance from center of column to neutral axis of transformed area considering bending only.

x' = center of gravity distance of A'_s from centerline of column.

$\lambda = \phi/180 = A'_s/A_s$ = ratio of additional compression steel to total area of ring where $\cos \phi = 1 - a/R'$ and $a = x - t'$.

References

(1) American Association of State Highway Officials
Standard Specifications for Highway Bridges. Seventh edition, Washington, D.C., 1957, pp. 92–97.

(2) Boguslavsky, Boris W.
Design of Reinforced Concrete. MacMillan Co., New York, 1956, pp. 139–149.

(3) Peabody, Dean, Jr.
The Design of Reinforced Concrete Structures. Second edition, Wiley & Sons, New York, 1946, pp. 255–289.

(4) Dunham, Clarence W.
The Theory and Practice of Reinforced Concrete. Second edition, McGraw-Hill, New York, 1944, pp. 197–204.

(5) Large, George E.
Basic Reinforced Concrete Design: Elastic and Creep. Second edition, Ronald Press, New York, 1957, pp. 117–119.

(6) Graham, John W.
Discussion of *Reinforced Concrete Columns Under Combined Compression and Bending.* Proceedings of the American Concrete Institute, vol. 18, No. 10, June 1947, p. 8–5.

Sampling Techniques Applicable to the Collection of Economic Data

BY THE HIGHWAY NEEDS AND ECONOMY DIVISION
BUREAU OF PUBLIC ROADS

Reported by NATHAN LIEDER,
Statistician

Sample survey techniques are used in many data-collecting phases of highway research. However, not all who must practice sampling have been trained in modern sampling theory. This article presents an account of the theory and its application which is directed primarily to those engaged in sampling for economic data.

Included is a discussion of the basic concepts of sampling, and the advantages and disadvantages of several alternative sampling methods. Proportional stratified sampling and optimum allocation among strata are discussed. Criteria are offered for choosing between these two alternatives with some consideration given to cost factors.

The author suggests one possible application of the theory and offers an illustration based on hypothetical data.

ADMINISTRATORS require sufficient and accurate data upon which to base decisions on broad general policies, as well as on day-to-day problems. Decisions must be made whether pertinent data are available or not. Therefore, data collection and data analyses are the ever-present concern of administrators.

Solution of the problem of data collection is made more difficult by several limiting factors. Two of these are lack of time and lack of money. Sampling is a tool to minimize, to some extent, the effect of these limiting factors. Because sampling cannot effect a limitless spread of resources, administrators must decide on priorities among the data to be collected and on the amount of money to be spent in collection and analysis.

Modern statistical theory not only lights the way to good sample design as an aid to data collection, but it also emphasizes techniques for drawing logical inferences from the the data as an aid to analysis. This latter aspect of statistical theory is not sufficiently emphasized or practiced.

Sampling is not just a stop-gap or substitute for the collection of all data relating to a particular study. It is often not realized that estimates based on relatively small samples may be more reliable than results of complete censuses. Since fewer personnel are needed for sample studies than for complete censuses, there is opportunity for more thorough training and for better supervisory control of employees than is possible in complete censuses. Because data to be processed in a limited time are not amassed for sample studies in the same huge quantities as for complete censuses, an opportunity is presented for better control of the various data processing operations to attain better quality of output. The stress is laid on the word "opportunity." Sampling does not assure quality; it makes improved quality possible with the funds available.

All decisions based on the data collected by the State highway departments and by the Bureau of Public Roads are supported, in one sense, by sample studies. Even complete censuses may be considered sample studies; they are samples in time. Administrators must assume a certain degree of stability or variability over time when making decisions based on such studies. Therefore, anyone who decides to use a complete count is not voting against sampling, but rather against a sample of a given size. In other words, the individual believes that the estimates yielded by the smaller sample will not be sufficiently reliable for the decisions that must emanate therefrom.

Distribution of Sample Estimates

Statistical theory is based on probability samples only. Probability samples are those in which the chance of any element being selected can be calculated before selection and can be expressed as a mathematical ratio. When this condition is satisfied, the probability calculus can be applied to the sample data to make possible meaningful deductions on the reliability of the results, and to fashion improved designs for future sample studies. Use of the probability calculus is impossible with non-probability samples.

Samples yield estimates of the true values of the characteristics of a population. In this context, the word "population" represents the totality of all elements from which a sample may be drawn; the true value is obtained from a study of all elements of a population, not a sample. The word "true" does not imply that the value is fixed even for a point in time. The methods used in determining a true value also influence its magnitude. For example, interviewers possess a given competency and training. The designs of questionnaires and other forms have varying efficiencies in extracting data of high quality. Supervisors and data processors have their own degrees of competency. A change in any of these components will likely result in a different true value.

Sample data collected and processed under a given set of conditions will yield an estimate of the true value that would have been obtained from all elements under the same set of operating conditions. The sample estimate will differ from the true value by an amount of unknown magnitude. Another sample of the same size as the first will yield another estimate of the true value. Usually the estimate yielded by the second sample will differ from the one yielded by the first sample.

Standard deviation

Consider the estimates from all possible samples of the same size drawn from a specified population. These estimates from all possible samples of the same size form another population of their own—a population of sample estimates. They will have a distribution from low estimate to high estimate. As a general rule, the estimates will tend to cluster about the true value. The measure of the scatter of the population of sample estimates about their average value, generally the true value, is called the standard deviation of the population. The standard deviation of the derived population is a function of the standard deviation of the original population.

Normal distribution

The distribution of the population of sample estimates tends toward the normal. The graph of the mathematical function describing the normal distribution exhibits a symmetrical, bell-shaped curve. The area under this curve is distributed in a fixed fashion, whatever the values assigned to the constants in the mathematical function. Distributions of many types of sample estimates approximate the normal with increasing sample size. The larger the sample size, the more closely will the derived population of sample estimates approach the normal distribution. Furthermore, the larger the sample size, the more closely will the sample estimates cluster around the true value, and consequently the smaller will be the standard deviation. The normal distribution is a very satisfactory approximation of the distribution of estimates derived from samples of the sizes ordinarily used by highway departments for economic studies.

Hypothetical population

Some of the remarks on the distribution of sample estimates may be clarified by an illustration based on a hypothetical population of six households which compose a hypothetical State. Assume that residents of five of these households own passenger cars and travel the following annual mileages:

Household number	*Annual mileages*
1	0
2	2,000
3	6,000
4	7,000
5	12,000
6	18,000
Total travel	45,000

Further assume that it is desired to select a simple random sample of these households in order to estimate the total vehicle-miles for the State, and that the estimate is obtained by multiplying the sample values by the reciprocal of the sampling fraction.

The distribution of estimates from all possible samples, as well as the value of the standard deviation and the square of the standard deviation (the variance) for each sample size, is reported in table 1. The formula for calculating the variance, σ^2, is,

$$\sigma^2=\frac{\sum_{i}^{N} X_i^2}{N}-\frac{\left(\sum_{i}^{N} X_i\right)^2}{N^2}$$

Where:

N = The total number of possible samples.

X_i = The estimate obtained from the i^{th} sample.

The average values for travel shown immediately below the column totals of table 1 illustrate what is meant by an unbiased estimate. A technique yields an unbiased estimate when the average value of all possible estimates of the same item obtained by the application of the technique equals the population value that is being estimated. In this example, the population value is 45,000 vehicle-miles of travel.

A frequency distribution of the estimates in table 1 is presented in table 2 to emphasize the tendency of sample estimates to cluster in a fairly symmetrical fashion about a central value with increasing sample size.

Table 1.—All possible sample estimates of total travel and all corresponding standard deviations derived from a hypothetical population of 6 households

Sample number	Estimated travel (in thousands of vehicle-miles), based on all possible samples of—					
	1 household	2 households	3 households	4 households	5 households	All households
1	0	6	16	22.5	32.4	45
2	12	18	18	30.0	39.6	
3	36	21	26	31.5	45.6	
4	42	24	28	37.5	46.8	
5	72	27	30	39.0	51.6	
6	108	36	36	40.5	54.0	
7		39	38	40.5		
8		42	40	46.5		
9		54	40	48.0		
10		54	42	49.5		
11		57	48	54.0		
12		60	50	55.5		
13		72	50	57.0		
14		75	52	58.5		
15		90	54	64.5		
16			60			
17			62			
18			64			
19			72			
20			74			
Total, all samples	270	675	900	675	270	45
Average per sample	45	45	45	45	45	45
Variance	1,317.0	526.8	263.4	131.7	52.68	0
Standard deviation	36.3	23.0	16.2	11.5	7.3	0

Standard Error of Estimate

Ordinarily, the results of only one sample are available, not of all possible samples. From this one probability sample can be calculated not only an estimate of the true value, but also an estimate of the standard deviation of the population of sample true values. The latter estimate is called the standard error of estimate. The larger the sample size for a given design, the smaller will be the standard error of estimate.

Using the sample estimate of the true value and the standard error of estimate calculated from the sample, a range of values can be determined within which the unknown true value probably lies. This statement may or may not be correct in a given situation. Because the distribution of estimates from large samples tends to approximate the normal, it can be shown that, were similar statements made for all possible samples of the same size, using the individual sample estimates each time, approximately 67 percent of the statements would be correct, provided that the range were based on one standard error of estimate. Approximately 95 percent of the statements would be correct, provided the range were based on two standard errors of estimate. In any given case, it may be assumed that the sample is one of 67 out of 100 (one standard error) that yields ranges which include the true value; or, that it is one of 95 out of 100 that yields such ranges based on two standard errors. The confidence levels based on the results of a single sample can therefore be controlled by the number of standard errors that are used in defining the range. Moreover, for a given confidence level, the larger the sample size, the narrower is the range within which the true value probably lies.

Within practical limits, the more important the decision to be based on the results of a sample, the more money will be paid to reduce the range, determined for a given confidence level, within which the true value lies. In other words, the more important the use to which the estimate will be put, the greater is the precision desired in the estimate.

Some of the characteristics of standard errors of estimate and of their squares, the variances, can be illustrated with the hypothetical population of six households. The values for the variances and standard deviations, shown in table 3 and illustrated in figure 1, were derived from the formula,

$$s^2=\left[N^2\frac{N-n}{Nn}\right]\left[n\sum_{i}^{n} x_i^2-\left(\sum_{i}^{n} x_i\right)^2\right]\div\frac{1}{n(n-1)}$$

Where:

N = The population (six households).

n = The sample size.

x_i = The value of the i^{th} household in a sample.

The following observations may be made concerning table 3: (1) A sample containing only one unit cannot yield an estimate of the spread of the population of estimates from all possible samples which contain only one unit. (2) The formula results in an unbiased estimate of the variance as shown by comparing the values on the last line of table 3 with the next to the last line of table 1. (3) Similarly, the standard error of estimate is not an unbiased estimate of the standard deviation of the population of estimates from all possible samples of a given size. The bias decreases with increases in sample sizes. (The sample sizes ordinarily employed by the State highway departments in planning studies should result in standard errors with negligible bias.) (4) The larger the sample size, the more stable

Table 2.—Frequency distribution of all possible sample estimates of total travel derived from a hypothetical population of 6 households

Class limits of estimated travel (1,000 vehicle-miles)	Number of samples having indicated range of travel, based on all possible samples of—					
	1 household	2 households	3 households	4 households	5 households	All households
0– 9.9	1	1				
10–19.9	1	1	2			
20–29.9		3	2	1		
30–39.9	1	2	3	4	2	
40–49.9	1	1	4	5	2	1
50–59.9		3	4	4	2	
60–69.9		1	3	1		
70–79.9	1	2	2			
80–89.9						
90–99.9		1				
100–109.9	1					

Table 3.—Standard errors of estimates, *s*, and their squares, s^2, for each of all possible estimates of travel derived from a hypothetical population of 6 households

Sample number	Standard errors and their squares for estimates of travel, based on all possible samples of—											
	1 household		2 households		3 households		4 households		5 households		All households	
	s^2	s	s^2	s	s^2	s	s^2	s	s^2	s	s^2	s
1	(1)	(1)	24	4.9	56	7.5	32.75	5.7	26.16	5.1	0	0
2	(1)	(1)	216	14.7	78	8.8	84.00	9.2	58.56	7.7	-------	-------
3	(1)	(1)	294	17.1	86	9.3	86.75	9.3	65.76	8.1	-------	-------
4	(1)	(1)	96	9.8	248	15.7	72.75	8.5	65.04	8.1	-------	-------
5	(1)	(1)	150	12.2	42	6.5	195.00	14.0	54.96	7.4	-------	-------
6	(1)	(1)	864	29.4	216	14.7	194.75	14.0	45.60	6.8	-------	-------
7	-------	-------	6	2.4	218	14.8	50.75	7.1	-------	-------	-------	-------
8	-------	-------	600	24.5	584	24.2	168.75	13.0	-------	-------	-------	-------
9	-------	-------	1,944	44.1	152	12.3	216.00	14.7	-------	-------	-------	-------
10	-------	-------	216	14.7	150	12.2	140.75	11.9	-------	-------	-------	-------
11	-------	-------	150	12.2	504	22.4	180.00	13.4	-------	-------	-------	-------
12	-------	-------	1,536	39.2	494	22.2	174.75	13.2	-------	-------	-------	-------
13	-------	-------	864	29.4	62	7.9	147.00	12.1	-------	-------	-------	-------
14	-------	-------	726	26.9	416	20.4	140.75	11.9	-------	-------	-------	-------
15	-------	-------	216	14.7	402	20.0	90.75	9.5	-------	-------	-------	-------
16	-------	-------	-------	-------	504	22.4	-------	-------	-------	-------	-------	-------
17	-------	-------	-------	-------	266	16.3	-------	-------	-------	-------	-------	-------
18	-------	-------	-------	-------	392	19.8	-------	-------	-------	-------	-------	-------
19	-------	-------	-------	-------	216	14.7	-------	-------	-------	-------	-------	-------
20	-------	-------	-------	-------	182	13.5	-------	-------	-------	-------	-------	-------
Total, all samples	-------	-------	7,902	296.2	5,268	305.6	1,975.50	167.5	316.08	43.2	0	0
Average per sample	-------	-------	526.8	19.7	263.4	15.3	131.70	11.2	52.68	7.2	0	0

[1] Not applicable.

is the value of the standard error from sample to sample.

The length of each horizontal line in figure 1 represents twice the standard error given in table 3 for each of the sample estimates. The starting point, indicated by the fork at one end of each horizontal line, is the estimate of total travel shown in table 1 for each sample of a given sample size. Each horizontal line is extended in the direction of the true value of total travel, the heavy vertical line at the scale value of 45. If the sample estimate is less than 45, the value of 2 standard errors is added; and if the sample estimate is greater than 45, the value of 2 standard errors is subtracted. The spacing between horizontal lines has no significance.

Figure 1 shows that the various sample estimates of total travel (table 1) plus or minus twice the corresponding standard errors (table 3) produce ranges of values which include the population total that is being estimated in the following proportion of cases: sample size 2, 11/15; sample size 3, 15/20; sample size 4, 13/15; and sample size 5, 5/6. These proportions are poor approximations of the proportion of 95 in 100 which was previously mentioned. It must be remembered that the illustration is based on a small population with small sample sizes.

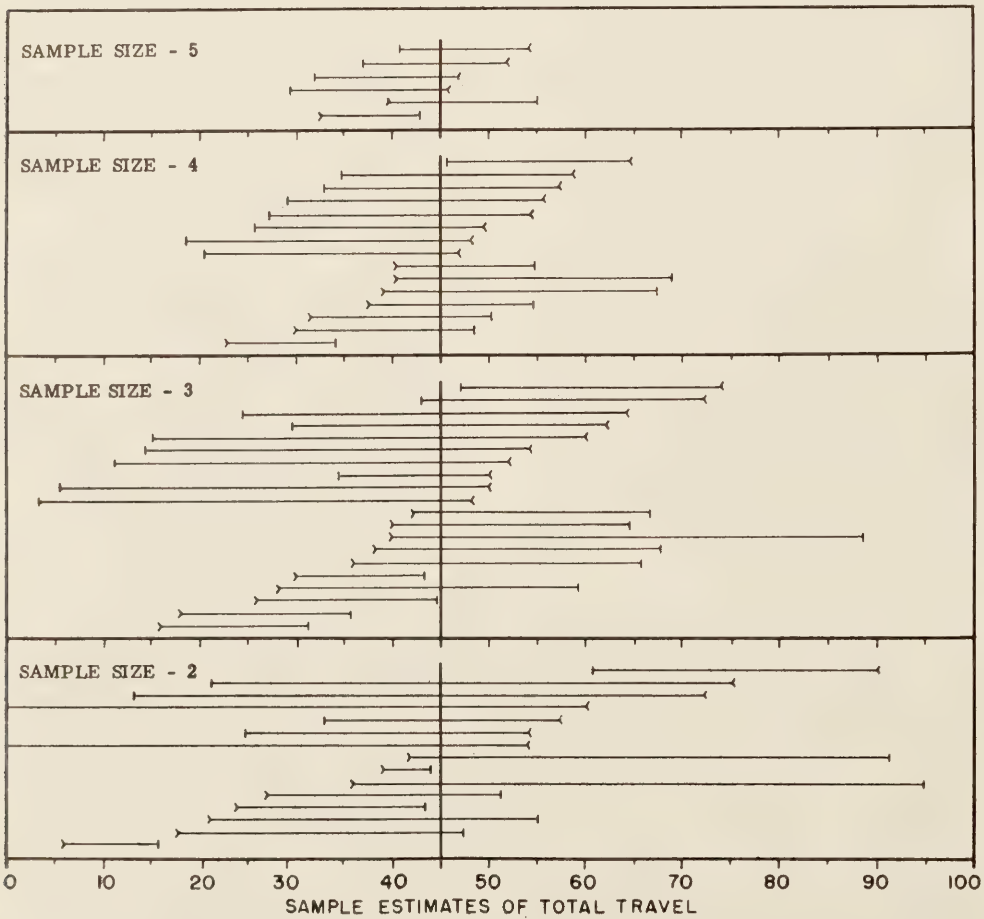


Figure 1.—All possible samples of a given sample size, plus or minus twice the corresponding standard errors (45 is the population value being estimated).

Optimum Sample Design

An important area in modern sampling theory has been the development of criteria for designing samples that will yield optimum results. The optimum is defined either as a specified standard error for the least cost, or as the smallest standard error for a given cost. Two sets of estimates and a predetermined standard are required in order to make use of the theory. The required sets are estimates of the unit costs of various study operations and the standard deviations of the several groupings into which will be allotted all the elements making up the original population. The organization paying for the study must predetermine either the total amount of money to be spent on a study, or the standard error to be attained for a given sample estimate. Often the latter is specified and an upper limit is placed on the money to be spent.

Estimates of the standard deviations of the several groupings must be calculated from data obtained in previous samples, from small-scale pilot studies, or from more or less educated guesses. Estimates of unit costs can be based on past experience.

Insufficient data on variances and costs have made it difficult to ascertain whether an optimum sample design has been attained in past studies in which the Bureau of Public Roads has been interested. However, efforts have been made to achieve a specified standard error. For example, in both the latest motor-vehicle-use studies and motor-vehicle accident cost studies, the samples were designed to attain standard errors that would be about 7 percent of the estimate for a given total at the level of one standard error.

One complication in the consideration of optimum design is that economic and traffic studies have been multi-purpose studies, and will likely continue to be multi-purpose studies. It is impossible to design one sample survey to produce a multitude of estimates, each with its own prespecified standard error. Someone must determine which single estimate, or possibly two or three estimates, are the most important. The sample is designed to attain the specified standard error or standard errors for these important estimates. The standard errors of the other subsidiary estimates must be accepted at whatever value the sample yields. For example, in the motor-vehicle-use studies the aim is now to attain the specified standard errors for two estimates: the total vehicle-miles for passenger cars driven by residents of a State, and the total vehicle-miles for trucks registered in a State. In the accident cost studies for passenger cars, the aim is now to attain the specified standard errors for three estimates: direct costs of fatal accidents, injury accidents, and property-damage-only accidents.

Considerations in Sample Design

Basic elements

Thus far, some of the broad aspects of sample design have been viewed from afar. At this point some of the details of the problem of sample design will be considered at close range. This consideration begins with two definitions: An *elementary unit* is an individual member of the particular population on which measurements are desired and for which analyses are to be made from the survey results. A *sampling unit* is either one elementary unit or a number of elementary units which are grouped and treated as one for sampling purposes.

Equating precision and cost

Determination of the optimum sampling unit is not always simple. For example, a city may be regarded as a number of city blocks, a number of households, or a number of persons. The optimum unit is one which gives the desired precision for the sample estimate at the smallest cost, or the greatest precision for a fixed cost. For a given sample size, a large sampling unit is usually less expensive than a small unit, but it is often less precise. To attain a specified precision in the overall sample, it is often necessary to get data from more elementary units when they are grouped into sampling units than if the elementary units are sampled directly. The choice of sampling unit involves striking a balance between relative precision and relative cost. Elementary units are grouped most often into sampling units when the sample design calls for one or more stages of subsampling.

Two types of grouping of sampling units are practiced: units grouped into strata and units grouped into clusters. Each type is formed under different conditions, but the underlying sampling principle is to achieve maximum precision per unit of cost. Both types of groupings may be practiced in the same sample design; use of one does not eliminate the other.

Stratified samples

A stratified sample is one in which the elements of the population are separated into fairly homogeneous groups and a sample is drawn to represent each group, each sample being independent of the one drawn from any other group. Stratification is resorted to in the hope of attaining increased precision for a given sample size, or a specified precision for a smaller sample size. The amount of increase in precision of sample estimates accomplished by stratification will depend on the degree of homogeneity that is achieved within strata. This, in turn, depends on how effectively the strata have been defined.

In defining strata, use should be made of all information that helps classify members of the population into groups which differ from one another with respect to the characteristic being measured, or with respect to the cost of collecting data. But, within each stratum, the sample must be a probability sample.

Unless relatively large differences can be set up among the strata, the gain in precision will be only moderate or perhaps barely noticeable. Stratification is sometimes adopted for administrative convenience and money savings with no prospect of reducing standard errors. This often occurs when a stable field organization exists, and it is less costly to use the established setup than to form a temporary organization for a single effort.

The criteria used in forming strata should be fairly well correlated with the changes in the value of the important characteristic so that stratum boundaries may be readily determined. As a general rule, most of the gain in precision has been attained with the use of one criterion that is strongly correlated with the values of the item to be estimated. Other criteria that are well correlated with the values of the item to be estimated are probably also well correlated with the first criterion. For two reasons, therefore, it is often uneconomical to spend much time on stratification. In the first place, stratification may provide little gain in precision since no marked differences can be segregated; and secondly, most of the gain has been realized in the application of one or two well-correlated criteria.

The preceding discussion should not lead anyone to forget the obverse. Stratification can yield marked gains in precision, the amount of the gain being proportional to sizes of the differences among the strata.

A general rule for optimum results is that the number of sampling units to be selected from any stratum should be directly proportional to the variability of the units and inversely proportional to the square root of the costs for collecting and processing the data from the stratum.

Cluster samples

The formation of clusters for sampling purposes is most often resorted to in area sampling. Clusters are formed from the sampling units or elementary units that are found in a more or less compact area. Cluster sampling may also be applied to lists or to files of reports as a technique to save examination of every unit or to save listing costs.

The guiding principle in the formation of clusters is that the component units be as heterogeneous as possible. This is the opposite of the rule for forming strata. Insofar as possible, each cluster is made as heterogeneous as the population, or as some subpopulation. If this were perfectly possible, the examination of all elements in one cluster would be sufficient to yield an acceptable estimate. This is generally not possible. Contiguous units tend to have similar characteristics. Because of this tendency, a cluster sample of a given size yields less precise estimates than stratified samples of the same size. Hence, somewhat larger sample sizes are required for cluster samples than for stratified samples to attain a desired precision of estimate. The increase in sample size is a function of the average similarity of the units composing the clusters.

Despite the increase in sample size, cluster sampling often results in reduced costs, and the cost reduction may be substantially greater than that experienced with simple random sampling of ungrouped data or stratified sampling. These cost economies are realized in savings in travel costs, administrative costs, and listing costs.

Stratification and cluster sampling may be used in the same sample survey. They are merely methods of grouping. Stratification can yield a gain in precision for a given sample size or result in a smaller sample for a given precision. Clustering can yield a savings in certain aspects of survey costs despite an increase in sample size.

Selecting the Sampling Method

Whether sampling units are grouped or not, they must still be selected for a study. Three selection methods are in common use: simple random sampling, systematic sampling, and sampling with unequal probabilities. Since the first two methods are often treated synonymously, and since the third may be unfamiliar, it may be well to define these terms.

A *simple random sampling* of n sampling units is one selected so that each combination of the n units has the same chance of being selected as any other combination.

A *systematic sample* of n sampling units is one in which a constant selection interval is applied n times in step fashion to the ordered elements of a population.

A *sample selected with unequal probabilities* is one in which the sampling units composing a population have been assigned varying weights, and the chance for the selection of any unit is proportional to its assigned weight.

The basic difference between systematic and simple random samples is that not every com-

(*Continued on page 252*)

528768—59——2

Prepared by th

Line	State	Width inches[1]	Height ft.-in.	Length—feet[2]: Single unit: Truck	Single unit: Bus	Truck tractor semi-trailer	Other combi-nation	Number of towed units[3]: Semi-trailer	Full trailer	Semi-trailer and full trailer	Axle load—pounds: Single: Statutory limit	Single: Including statutory enforcement tolerance	Tandem: Statutory limit	Tandem: Including statutory enforcement tolerance	Type of restrict
1	Alabama	96	[6] 12-6	35	40	50	NP	1	NP	NP	18,000	19,800	36,000	39,600	Table
2	Arizona	96	13-6	40	40	65	65	1	1	2	18,000		32,000		Table
3	Arkansas	96	13-6	35	40	50	50	1	1	NP	18,000	[7] 18,500	32,000	32,500	Spec. maximum[8]
4	California	96	13-6	35	[9] 35	[10] 65	[10] 65	NR	NR	NR	18,000		32,000		Table
5	Colorado	[11] 96	[12] 13-6	35	40	60	60	1	2	2	18,000		36,000		Formula-spec. li
6	Connecticut	102	12-6	50	50	50	NP	1	NP	NP	22,400	22,848	36,000	36,720	Spec. lim.-tire ca
7	Delaware	96	[6] 12-6	40	42	50	60	1	1	2	20,000		36,000		Table-spec. limit[8]
8	District of Columbia	96	12-6	35	35	50	80	1	1	NP	22,000		38,000		Table
9	Florida	96	[6] 12-6	[14] 35	40	50	50	1	1	NP	20,000	22,000	40,000	44,000	Table
10	Georgia	96	13-6	[15] +39	[15] +45	50	50	1	1	NP	18,000	20,340	36,000	40,680	Spec. maximum[16]
11	Hawaii	108	13-0	40	40	55	65	1	1	2	24,000		32,000		Formula[17]
12	Idaho	[18] 96	14-0	35	[19] 40	60	65	1	1	2	[20] 18,000		[20] 32,000		Table[20]
13	Illinois	96	13-6	42	42	50	50	1	1	2	[21] 18,000		32,000		Spec. lim.-tire ca
14	Indiana	96	13-6	36	40	50	50	1	1	2	[23] 18,000	[23] 19,000	[23] 32,000	[23] 33,000	Spec. lim.-tire c
15	Iowa	96	13-6	35	[19] 40	50	NP	1	[24] 1	NP	18,000	18,540	32,000	32,960	Table
16	Kansas	96	13-6	35	[19] 40	50	50	1	1	NP	18,000		32,000		Table
17	Kentucky	96	[25] 13-6	[26] 35	[26] 35	[27] 50	NP	1	NP	NP	18,000	[28] 18,900	32,000	[28] 33,600	Spec. lim.-tire c[29]
18	Louisiana	96	[6] 12-6	35	[19] 40	50	60	1	1	NP	18,000		32,000		Axle lim.-tire ca
19	Maine	96	[30] 12-6	50	50	50	50	1	1	NP	[30] 22,000		[30] 32,000		Table-tire cap.
20	Maryland	96	[6,45] 12-6	55	55	55	55	NR	NR	NR	22,400		[31] 40,000		Formula
21	Massachusetts	96	NR	35	[19] 40	[31] 50	NP	1	NP	NP	22,400		36,000		Table-spec. lim
22	Michigan	96	13-6	35	40	55	55	1	1	2	[33] 18,000		[34] 32,000		Axle lim.-tire c.
23	Minnesota	96	13-6	40	40	50	50	1	1	NP	18,000		32,000		Table
24	Mississippi	96	[6] 12-6	35	40	[36] 50	50	1	1	NP	18,000		28,650	[35] 32,000	Table-tire cap.
25	Missouri	96	12-6	35	40	50	50	1	1	2	18,000		32,000		Table
26	Montana	[18] 96	13-6	35	40	60	60	1	1	[37] 2	18,000		32,000		Table
27	Nebraska	96	13-6	40	40	60	60	1	1	2	18,000	18,900	32,000	33,600	Table
28	Nevada	96	NR	NR	NR	NR	NR	NR	NR	NR	18,000	18,900	32,000	33,600	Table
29	New Hampshire	96	13-6	35	[35] 40	50	50	NR	NR	NR	22,400		36,000		Tables-spec. li
30	New Jersey	96	13-6	35	[39] 35	45	[40] 50	1	1	NP	22,400	23,520	32,000	33,600	Spec. limits
31	New Mexico	[41] 96	13-6	40	40	65	65	1	1	2	21,600		34,320		Table
32	New York	96	13-0	35	[42] 35	50	50	1	1	NP	22,400		36,000		Formula
33	North Carolina	96	[6] 12-6	35	[19] 40	[43] 50	[43] 50	1	1	NP	18,000	19,000	36,000	38,000	Spec. limits
34	North Dakota	96	13-6	[14] 35	[19] 40	60	60	1	1	2	18,000		32,000		Formula
35	Ohio	96	13-6	35	[19] 40	50	60	1	NR	NR	19,000		[44] 31,500		Formula
36	Oklahoma	96	13-6	35	45	50	50	1	1	NP	18,000		32,000		Table
37	Oregon	96	[45] 12-6	35	[35] 40	[46,35] 55	[35] 65	1	1	[35] 2	[47] 18,000		[47] 32,000		Table[48]
38	Pennsylvania	96	[6] 12-6	35	[19] 40	50	[40] 50	1	1	NP	22,400	23,072	36,000	37,080	Spec. limits[49]
39	Puerto Rico	96	12-6	NS	NS	NS		NS	NS	NS	NS		NS		Spec. lim.-tire
40	Rhode Island	102	12-6	40	40	50	50	1	1	NP	22,400		NS		Spec. limits
41	South Carolina	96	[45] 12-6	[14] 35	[19] 40	50	60	1	1	NP	20,000		32,000		Table
42	South Dakota	96	13-0	35	40	50	60	1	1	2	18,000		32,000		Table
43	Tennessee	96	[6] 12-6	35	40	50	50	1	[53] 1	NP	18,000		32,000		Table
44	Texas	96	13-6	35	40	50	50	1	1	NP	18,000	18,900	32,000	33,600	Table[54]
45	Utah	96	14-0	45	45	60	60	NR	NR	NR	18,000		33,000		Table
46	Vermont	96	12-6	50	50	50	50	1	1	NP	NS		NS		Spec. lim.-tire c
47	Virginia	96	[6] 12-6	35	[35] 40	50	50	1	1	NP	18,000		[57] 32,000		Table
48	Washington	96	13-6	35	[19] 40	60	[58] 65	1	1	[58] 2	18,000	[59] 18,500	32,000	[59] 33,000	Table-spec. lim
49	West Virginia	96	[6] 12-6	35	[19] 40	50	50	1	1	NP	18,000	18,900	32,000	33,600	Table
50	Wisconsin	96	[6] 12-6	35	40	50	50	1	1	NP	18,000	[60] 19,500	30,000	32,000	Table-formula[61]
51	Wyoming	96	13-6	40	40	60	60	1	1	2	18,000		32,000	[62] 36,000	Table
52	Alaska	96	12-6	35	[19] 40	60	60	1	1	2	18,000		32,000		Table-tire cap.
	AASHO Policy	96	12-6	35	[19] 40	50	60	1	1	NP	18,000		32,000		Table
	Number of States: Higher	3	30	18	29	18	10	6	8	25	31		29		Formula 6
	Number of States: Same	49	22	34	19	33	12	46	40	27	21		21		Table 2
	Number of States: Lower	0	0	0	4	1	30	0	4	0	0		2		Specified limits 4

NP—Not permitted. NR—Not restricted. NS—Not specified.

[1] Various exceptions for farm and construction equipment; public utility vehicles; house trailers; urban, suburban, and school buses; haulage of agricultural and forest products; at wheels of vehicles; for safety accessories, on designated highways, and as administratively authorized.

[2] Various exceptions for utility vehicles and loads, house trailers and mobile homes.

[3] When not specified, limited to number possible in practical combinations within permitted length limits; various exceptions for farm tractors, mobile homes, etc.

[4] Legally specified or established by administrative regulation.

[5] Computed under the following conditions to permit comparison on a uniform basis between States with different types of regulation:

A. Front axle load of 8,000 pounds.

B. Maximum practical wheelbase within applicable length limits:

(1) Minimum front overhang of 3 feet.

(2) In the case of a 4-axle truck-tractor semitrailer, rear overhang computed as necessary to distribute the maximum possible uniform load on the maximum permitted length of semitrailer to the single drive-axle of the tractor and to the tandem axles of the semitrailer, within the permitted load limits of each.

(3) In the case of a combination having 5 or more axles, minimum possible combined front and rear overhang assumed to be 5 feet, with maximum practical load on maximum permitted length of semitrailer, subject to control of loading on axle groups and on total wheelbase as applicable.

C. Including statutory enforcement tolerances as applicable.

[6] Auto transports 13 feet 6 inches.

[7] Does not apply to combinations of adjacent load-carrying single axles.

[8] 56,000 pounds on load-carrying axles, exclusive of steering-axle load.

[9] On specific routes in urban or suburban service under special permit from P.U.C. 40 feet, also 3-axle buses with turning radius less than 45 feet without restriction.

[10] Effective September 18, 1959 on Interstate, 4-lane, and designated State highways.

[11] Buses 102 inches.

[12] On designated highways; 12 feet 6 inches on other highways.

[13] Legal limit 60,000 pounds, axle spacing 27 feet or more.

[14] Three-axle vehicles 40 feet.

[15] Truck 39.55 feet; bus 45.20 feet.

[16] 63,280 pounds maximum, except on roads under R

[17] 700 (L+40) when L is 18' or less; 800 (L+40) wh with span of 20' or over.

[18] Buses 102 inches on highways of surfaced width

[19] Less than three axles 35 feet.

[20] Special limits for vehicles hauling timber and tim including livestock; single axle 18,900 pounds, tandem mitted 66,000 pounds maximum at 21-foot axle spacing, foot axle spacing.

[21] On designated highways; 16,000 pounds on other

[22] Without tandem axles 45,000 pounds.

[23] On designated highways; single axle 22,400 poun excesses of weight under one or more limitations of axl

[24] Limited to 4 wheel trailer towed by truck not exce

[25] Class AA highways; 12 feet 6 inches on other hig

[26] On designated highways; trucks 26.5 feet and bus

[27] Class AA highways; 45 feet on other highways.

[28] Class AA highways only.

[29] Maximum gross weight on Class A highways 42,0

[30] Including load 14 feet; various exceptions for veh

[31] Effective September 21, 1959.

[32] Subject to axle and tabular limits.

[33] Single axle spaced less than 9 feet from nearest

[34] On designated highways only and limited to one ta

[35] On designated highways only.

[36] Auto transports permitted 50 feet.

[37] Semitrailer and semitrailer converted to full traile

ads, July 1, 1959

Applicable to: Any group of axles	Applicable to: Total wheel base only	Specified maximum gross weight—pounds[4]: Truck 2-axle	Truck 3-axle	Truck-tractor semitrailer 3-axle	Truck-tractor semitrailer 4-axle	Truck-tractor semitrailer 5-axle	Other combination	Practical maximum gross weight—pounds[5]: Truck 2-axle	Truck 3-axle	Truck-tractor semitrailer 3-axle	Truck-tractor semitrailer 4-axle	Truck-tractor semitrailer 5-axle	Other combination	Line
	X							27,800	47,600	47,600	60,010	64,650	NP	1
Under 18'	Over 18'							26,000	40,000	44,000	58,000	72,000	76,800	2
								26,500	40,500	45,000	59,000	65,000	65,000	3
Under 18'	Over 18'							26,000	40,000	44,000	58,000	72,000	76,000	4
	X	30,000	46,000					26,000	44,000	44,000	62,000	76,000	76,000	5
		32,000	50,000	50,000	60,000	60,000	NP	30,848	44,720	51,000	61,200	61,200	NP	6
X		30,000	46,000	48,000	60,000	60,000	60,000	28,000	48,000	48,000	56,350	60,000	60,000	7
X								30,000	46,000	52,000	58,450	61,490	64,650	8
	X							30,000	52,000	52,000	65,200	73,095	73,095	9
							63,280	28,340	48,680	48,680	63,280	63,280	63,280	10
X								32,000	38,800	56,000	62,800	69,600	78,000	11
X								26,000	40,000	44,000	58,000	72,000	76,800	12
		36,000	[22] 41,000	45,000	59,000	72,000	72,000	26,000	40,000	44,000	58,000	72,000	72,000	13
							72,000	27,000	41,000	45,000	59,000	[23] 73,000	[23] 73,000	14
X								26,540	40,960	45,080	59,500	73,280	NP	15
X								26,000	40,000	44,000	55,470	63,890	63,890	16
		36,000	50,000	54,000	59,640	59,640	NP	26,900	41,600	45,800	59,640	59,640	NP	17
								26,000	40,000	44,000	58,000	72,000	76,000	18
X		32,000	[30] 50,000	50,000	60,000	60,000	60,000	30,000	40,000	52,000	60,000	60,000	60,000	19
	X			65,000	65,000	65,000	65,000	30,400	48,000	52,800	65,000	65,000	65,000	20
	X	[32] 46,000	[32] 60,000	[32] 60,000	[32] 60,000	[32] 60,000	NP	30,400	44,000	52,800	60,000	60,000	NP	21
								26,000	[35] 40,000	44,000	[35] 58,000	[35] 66,000	[35] 102,000	22
X								26,000	40,000	44,000	58,000	68,000	72,500	23
X								26,000	36,650	44,000	54,650	[35] 58,000	[35] 58,000	24
X								26,000	40,000	44,000	55,470	64,650	64,650	25
Under 18'	Over 18'							26,000	40,000	44,000	58,000	72,000	76,000	26
X		36,000	54,000	54,000	71,146	71,146	71,146	26,780	41,200	45,320	59,740	73,280	73,280	27
Under 18'	Over 18'							26,900	41,600	45,800	60,000	74,000	76,800	28
	X	33,400	[38] 47,500	52,800	66,400			30,400	44,000	52,800	66,400	66,400	66,400	29
		30,000	40,000	60,000	60,000	60,000	60,000	31,500	41,600	55,040	63,000	63,000	63,000	30
Under 18'	Over 18'							29,600	42,320	51,200	63,920	76,640	86,400	31
	X					65,000	65,000	30,400	44,000	52,800	65,000	65,000	65,000	32
		31,500	46,200	46,200	65,100	65,100	65,100	27,000	46,000	46,000	65,100	65,100	65,100	33
Under 18'	Over 18'							26,000	38,000	44,000	56,000	64,000	64,000	34
	X							27,000	39,500	46,000	58,500	71,000	78,000	35
X								26,000	40,000	44,000	58,000	72,000	73,280	36
Under 18'	Over 18'					[48] 74,000	[48] 76,000	26,000	40,000	44,000	58,000	[48] 72,000	[48] 76,000	37
		33,000	47,000	50,000	60,000	60,000	62,000	31,072	45,080	51,500	61,800	61,800	63,860	38
														39
		[50] 36,000	[51] 44,000	[52] 50,000	[13] 60,000	60,000	88,000	30,400	44,000	50,000	60,000	60,000	88,000	40
X								28,000	40,000	48,000	60,000	66,839	71,115	41
X								26,000	40,000	44,000	58,000	72,000	73,280	42
	X							26,000	40,000	44,000	58,000	61,580	43,500	43
X								26,900	41,600	45,800	60,500	75,200	75,600	44
X								26,000	41,000	44,000	59,000	74,000	79,900	45
		30,000	[55] 50,000	50,000	[56] 60,000	[56] 60,000	[56] 60,000	30,000	[55] 50,000	50,000	60,000	60,000	60,000	46
X						[35] 56,800	[35] 56,800	26,000	40,000	44,000	56,800	56,800	56,800	47
Under 18'	Over 18'	28,000	36,000	46,000	60,000	68,000	72,000	26,000	36,000	44,000	60,000	68,000	72,000	48
X								26,900	41,600	45,800	57,844	63,840	63,840	49
X								27,500	40,000	47,000	59,500	68,000	68,000	50
X								26,000	44,000	44,000	62,000	73,950	73,950	51
Under 18'	Over 18'	29,000	43,000	47,000	61,000	75,000	75,000	26,000	40,000	44,000	58,000	72,000	75,000	52
X								26,000	40,000	44,000	55,470	61,490	71,900	
20	18							29	27	29	47	40	25	
								22	19	22	3	2	0	
								0	5	0	1	9	26	

000 pounds maximum.
900 (L +40) on highways having no structures

vise as administratively authorized.

entrates, aggregates, and agricultural products
ss weight table: vehicle with 3 or 4 axles per-
xles permitted 79,000 pounds maximum at 43-

ounds; tolerance of 1,000 pounds on total of all

ight.

vays.

ighways 30,000 pounds.
lucts and construction materials.

unds.
n; otherwise 26,000 pounds.

[38] Dual-drive axles; otherwise 40,000 pounds.
[39] Or as prescribed by P.U.C.
[40] Exception for poles, pilings, structural units, etc., permitted 70 feet.
[41] On designated highways 102 inches.
[42] Trackless trolleys and buses 7 passengers or more, P.S.C. certificate 40 feet.
[43] Including front and rear bumpers.
[44] Approved equipment 73,280 pounds.
[45] Certain types of vehicles and commodities under special permit on designated highways up to 13 feet 6 inches.
[46] 60 feet allowed truck tractor semitrailer on major routes designated by permit.
[47] Logging vehicles permitted 3-foot wheelbase tolerance, 19,000-pound single axle, 34,000-pound tandem axle.
[48] Governs gross weight permitted on highways designated by resolution of State highway commission or by permit.
[49] Single unit truck with 4 axles permitted 60,000 pounds.
[50] Axles spaced less than 6 feet 32,000 pounds; less than 12 feet 36,000 pounds; 12 feet or more gross weight governed by axle limit.
[51] Single vehicle with 3 or more axles spaced less than 16 feet 40,000 pounds; less than 20 feet 44,000 pounds; 20 feet or more governed by axle limit.
[52] Tractor semitrailer with 3 or more axles spaced less than 22 feet 46,000 pounds; not less than 27 feet 50,000 pounds.
[53] Limited to 3,500 pounds.
[54] Effective January 1, 1960.
[55] On Interstate routes 40,000 pounds.
[56] Tandem axles on trailer equipped with adequate brakes.
[57] Vehicles registered before July 1, 1956, permitted limits in effect January 1, 1956, for life of vehicle.
[58] Three-unit combinations on designated highways.
[59] Vehicles hauling logs permitted wheelbase and gross weight tolerances. Discretionary enforcement tolerances not included in computation of practical maximum gross weights.
[60] Axle load 21,000 pounds on 2-axle trucks hauling unmanufactured forest products.
[61] On Class A highways.
[62] Based on ruling of Attorney General.

(*Continued from page 249*)

bination of n sampling units has a chance of being selected when sampling systematically. The emphasis here is on the word "combination," because in both schemes each sampling unit has the same weight and the same chance of selection. The difference in the two techniques can easily be seen in the following example.

Assume a population of 6 sampling units, numbered 1, 2, 3, 4, 5, and 6, and that it is desired to select 2 units for the sample. With simple random sampling, 2 random numbers less than or equal to 6 would be found in a table of random numbers to determine the units selected. One of the following 15 possible combinations would be chosen:

1, 2	2, 3	3, 5
1, 3	2, 4	3, 6
1, 4	2, 5	4, 5
1, 5	2, 6	4, 6
1, 6	3, 4	5, 6

With systematic sampling, a random start less than or equal to 3 would be found in a table of random numbers and every third sampling unit thereafter would be selected. One of the following 3 combinations would be chosen: 1 and 4, 2 and 5, or 3 and 6. Note that each sampling unit in a systematic sample can appear in only one combination of units. Each such combination may be considered a cluster. One cluster is selected in systematic sampling.

Simple random sampling

Simple random samples are generally not drawn for large-scale sample surveys. It is time-consuming and difficult to administer and control the selection and the use of n random numbers in a large-scale sample survey. However, the use of this technique, when it is beneficial from the cost standpoint, makes possible the calculation of an unbiased estimate of the contribution by a population or a stratum to the sampling variances.

Systematic sampling

Systematic sampling has two possible disadvantages. The first occurs when a population contains periodicities or is subject to trend. Systematic samples may then be ineffective in revealing the variation in the data. For example, assume that it was desired to sample enlisted personnel living on army posts in order to get data on motor-vehicle use. Further, suppose that a listing of names by barracks were available, that the name of a master sergeant generally headed the listing for each barracks, that 50 men were quartered in each barracks, and that the sampling rate were 1:25 with a random start of 1. The systematic sample would contain an excess of master sergeants whose characteristics, insofar as motor-vehicle use is concerned, would probably differ markedly from personnel of the lesser grades. This example also illustrates one of the criteria for cluster sampling—the need for making each cluster as heterogeneous as possible. As noted earlier, a systematic sample may be regarded as a single selected cluster. When a population contains periodicities and the sampling rate coincides to some extent with the period, the resulting sample is unduly homogeneous. The second disadvantage to systematic samples, usually surmountable, is that calculation of an unbiased estimate of the sampling variance is impossible with most designs employing this technique.

Systematic sampling has two important advantages. First, it is easy to administer; and second, if the arrangement of the population is such that the sampling units are essentially stratified, a systematic sample will be equivalent to a proportionate stratified sample with whatever gains in precision may be yielded by stratified sampling. For example, a file of reports may be arranged by counties within State highway districts. Sampling systematically from the file results in a sample that is stratified geographically.

Sampling with unequal probabilities

Underlying the techniques of simple random sampling and systematic sampling is the assumption that each sampling unit has approximately as much to contribute as any other sampling unit; hence, all units should be assigned equal weights. Also underlying these techniques is the second assumption that any one possible combination of sampling units is as acceptable as any other possible combination. When either of these two assumptions is grossly unrealistic, sampling with unequal probabilities is employed.

When the first assumption is not applicable, the technique most frequently adopted is to select sampling units with probability proportionate to some measure of the size of each sampling unit. Size does not necessarily mean volume or area. It means a function of the value of some characteristic, such as a total or the square root of a total, that is of interest in a sample study, or a function of the value of some other nonstudy characteristic that is highly correlated with the one being investigated. Sampling with probability proportionate to size often results in precision attained at a cost that is satisfactorily close to that yielded by an optimum design.

Sampling with probability proportionate to size has been practiced in the motor-vehicle-use studies conducted by some States. These studies are based on samples of dwelling units selected after several stages of sampling. In the first stage, all areas and places in a State are stratified. The estimated number of dwelling units in each area or place composing a stratum determines the chance for the selection of the area or place.

Problems of Nonresponse and Quality Control

In addition to the foregoing general treatment of some of the problems of sample design and sampling practices, two problems will now be discussed that apply as much to complete censuses as to sample surveys. These problems are nonresponse and quality control.

The larger the rate of nonresponse in any study, the more questionable the results of the study when applied to the entire population. It is often comfortably assumed that the nonrespondents have the same distribution of characteristics as respondents. Too often this assumption is not justified. Inferences of greater validity concerning the entire study population will be drawn if the nonresponse rate is reduced to such a level that, whatever the distribution of characteristics in the nonresponse group, it would not seriously have changed the distribution as estimated from the data received from the respondents. Research studies form the basis for decisions of some importance or they would not be conducted. The decisions should be based on relatively accurate data. A rule of thumb that has recently been applied by the Bureau of Public Roads in sample studies designed to obtain economic data is that the nonresponse rate should not be greater than 8 percent. To attain this rate or better can be costly, but possibly less costly than the results of decisions based on inaccurate estimates.

In modern sampling practice, considerable stress is given to the control of the quality of the final product. Sampling theory as described in this article and as presented in the literature aims at getting estimates of so-called true values with a given precision at minimum cost. Poor quality in questionnaire design, interviewer training, supervision, data processing, or any combination of these will generally yield different estimates of true value and precision than would have been obtained had a better quality of work been sought. Is a true value based on work of poor quality an acceptable outcome for a research study? The answer is obvious.

No inference is intended with respect to the quality of any past sample studies. The intent is to foster a greater awareness of procedural pitfalls that may beset researchers in areas other than sample design.

Variations in Stratified Sampling Procedure

Since stratified sampling has been adopted in many designs used by the State highway planning organizations, a closer look at two variations of this procedure may prove helpful. These variations are proportionate stratified sampling, in which the same proportion of sample cases is selected from each stratum, and optimum allocation, in which the number of sample cases to be selected from each stratum, not the proportion, is determined so as to yield optimum results. In both variations simple random sampling is used to select sampling units from each stratum.

Symbols will be used to generalize the exposition of the two variations. Some deductions will be drawn from the generalized expressions, and some explanation will be given. No attempt will be made to prove the formulas.[1]

[1] For a more detailed explanation of these and of other aspects of sample design, the reader may refer to *Sample Survey Methods and Theory, Volume I, Methods and Applications,* by Morris H. Hansen, William N. Hurwitz, and William G. Madow, or to *Sampling Techniques,* by William G. Cochran. Both textbooks were published by John Wiley and Sons, Inc , New York, 1953.

et:

N_h = Total number of sampling units in h^{th} stratum.

n_h = Number of sampling units selected from h^{th} stratum.

L = Total number of strata.

$N = \sum_h^L N_h$ = Total number of sampling units in the population.

$n = \sum_h^L n_h$ = Total sample size.

X_{hi} = Value of a characteristic of i^{th} unit in h^{th} stratum.

$X_h = \sum_i^{n_h} X_{hi}$ = Value of a characteristic for all units selected from h^{th} stratum.

$X = \sum_h^L \sum_i^{n_h} X_{hi}$ = Value of a characteristic for all sample units.

$\overline{X}_h = \frac{X_h}{n_h} = \frac{\sum_i^{n_h} X_{hi}}{n_h}$ = Mean of the characteristic for the sample units in h^{th} stratum.

$X'_h = N_h \overline{X}_h = \frac{N_h}{n_h} \sum_i^{n_h} X_{hi}$

= Estimate of a total for the h^{th} stratum.

$X' = \sum_h^L X'_h = \sum_h^L \frac{N_h}{n_h} \sum_i^{n_h} X_{hi}$ = Estimate of a total for the population.

$\frac{N_h}{n_h}$ = Reciprocal of the sampling fraction used in the h^{th} stratum.

$$s_h^2 = \frac{\sum_i^{n_h} (X_{hi} - \overline{X}_h)^2}{n_h - 1}$$

= An estimate from the sample of the square of the standard deviation, (the variance), of the population of sampling units in the h^{th} stratum.

$$S_h^2 = \frac{\sum_i^{N_h} \left(X_{hi} - \frac{\sum_i^{N_h} X_{hi}}{N_h} \right)^2}{N_h - 1}$$

= The variance of the population of sampling units in the h^{th} stratum.

y_{hi} = An estimate of the value of the i^{th} unit in the h^{th} stratum obtained from sources other than the sample.

$$T_h^2 = \frac{N_h \sum_i^{N_h} (y_{hi})^2 - \left(\sum_i^{N_h} y_{hi} \right)^2}{N_h (N_h - 1)}$$

= An estimate of the variance, S_h^2, of the population of sampling units in the h^{th} stratum obtained from sources other than the sample.

$$s_{x'}^2 = \sum_h^L N_h^2 \frac{(N_h - n_h)}{N_h n_h} s_h^2$$

= The square of the standard error of estimate of X'.

Since $n_1/N_2 = n_2/N_2 \ . \ . \ . = n/N$ in proportionate stratified sampling, a simplified formula for this design would be

$$s_{x'}^2 = \frac{N-n}{n} \sum_h^L N_h s_h^2$$

In proportionate stratified sampling, once n is known, the various n_h values are determined from $n_h = N_h \ (n/N)$. In optimum allocation, the various n_h values are determined by one of the two following formulas:

$$n_h = \frac{N_h S_h}{\Sigma N_h S_h} n \quad \text{or} \quad n_h = \frac{N_h S_h \div \sqrt{C_h}}{\Sigma (N_h S_h \div \sqrt{C_h})} n$$

Where:

C_h = The variable cost from stratum to stratum for collecting and processing the data.

The first allocation formula should be used in all cases except when the variation in the cost factors among the strata is of the magnitude of 3 or more, since it is the square root of the costs that enters into the formula.

Since proportionate stratified sampling uses a constant sampling fraction, it is relatively simple to produce estimates of population characteristics from the sample. This advantage is worth preserving unless optimum allocation produces a marked gain in sample efficiency. The following computations will aid in arriving at a choice between the two alternatives.

$$\sigma_{s_h}^2 = \frac{1}{N} \sum_h^L N_h (S_h - \overline{S})^2$$

or

$$\sigma_{T_h}^2 = \frac{1}{N} \sum_h^L N_h (T_h - \overline{T})^2$$

Where:

$$\overline{S} = \frac{\sum_h^L S_h}{L} \quad \text{and} \quad \overline{T} = \frac{\sum_h^L T_h}{L}$$

$$V_{S_h}^2 = \frac{\sigma_{S_h}^2}{(\overline{S})^2} \quad \text{or} \quad \frac{\sigma_{T_h}^2}{(\overline{T})^2}$$

When $V_{S_h}^2$ or $V_{T_h}^2$ equals one-third, optimum allocation will yield a squared standard error of estimate that is about 90 percent as large as that yielded by proportionate stratified sampling.

When V_{Sh}^2 or $V_{T_h}^2$ equals 1, optimum allocation will yield a squared standard error of estimate that is about 50 percent as large as that yielded by proportionate stratified sampling.

Therefore, it would be good practice to use proportionate stratified sampling when $V_{S_h}^2$ is of the magnitude of one-third or smaller; otherwise use optimum allocation.

Implicit in the preceding formulas is the assumption that good estimates are available for S_h, the population values of each stratum's standard deviation. When good estimates of the S_h values are not available, but it is possible to determine a measure of size, Y_h, for each stratum, and it is known that the Y_h values are fairly stable over the years, then a good approximation to optimum allocation can be determined by

$$n_h = n \frac{Y_h}{\sum_h^L Y_h}$$

So-called "optimum allocation" may yield less precise results than proportionate stratified sampling if good estimates of S_h values are not available or the Y_h values are not suitable.

Thus far, some formulas have been presented for determining the allocation of the total sample to each of the strata. No formulas have been given to calculate the total size of sample to be allocated. The following formulas may be applied in the case of proportional stratified sampling:

$$n_o = \frac{N}{d^2} \sum_h^L N_h S_h^2 \quad \text{or} \quad \frac{N}{d^2} \sum_h^L N_h T_h^2$$

Where:

n_o = First estimate of total sample size.

d^2 = The square of the desired standard error of estimate.

If n_o/N is of the magnitude of one-tenth or greater, compute

$$n = \frac{n_o}{1 + \frac{n_o}{N}}$$

The following formulas may be used for optimum allocation:

$$n = \frac{\left(\sum_h^L N_h S_h \right)^2}{d^2 + \sum_h^L N_h S_h^2} \quad \text{or} \quad \frac{\left(\sum_h^L N_h T_h \right)^2}{d^2 + \sum_h^L N_h T_h^2}$$

In area sampling especially, strata often vary considerably in size. A good rule when different sampling fractions are used for the several strata, as in optimum allocation, is to define strata measuring approximately equal in size. It may be found that some areas or clusters of sampling units are much larger than the generality. Each of these large areas may be set apart as a self-representing area composing a stratum. A satisfactory arbitrary rule to follow is to define a stratum with only one area or extra large cluster when the following relationship holds.

$$Y_{area} = \frac{Y}{2n}$$

Where:

Y_{area} = Measure of size for the area.
Y = Total measure of size.
n = Desired sample size.

Possible Application of Sampling Procedures

The theory and formulas presented in this article should find ready application in much of the data collection activities of highway planning organizations. One such activity may be singled out as an illustration, namely, the collection of data on local government finance. In most States, data are collected from the entire population of local governments. A stratified sample may be designed as an alternative.

Local governments may be stratified for example on the basis of total receipts, total expenditures, or the sum of the two. If some local governments levy one type of tax which others do not, this may serve as a basis for stratification. The relatively few large cities may be set up as self-representing areas. Values of T_h^2 for each stratum may be computed from data for the preceding year. The N_h values will be known. Either a proportional stratified sample or a sample using optimum allocation may be drawn, depending on the value of $V^2_{T_h}$. If it is decided to adopt optimum allocation, the allocation formula probably need not take the square root of cost, $\sqrt{C_h}$, into consideration.

The opportunity to improve the quality of the data should not be overlooked. It may be possible, with sampling, to modify forms and techniques to assure better quality.

Modification of procedure

One objection that has been raised to the application of sampling in this field is that some important items of local finance, such as bond issues, are too rare and unpredictable for a sample to result in acceptable estimates. This is true. However, it is also true that it is uneconomical to include the whole population in a study in order to obtain an acceptable value for a scarce item. The alternative is to find some inexpensive method for singling out such items for special treatment. One possible means is to inquire by mail or telephone to determine whether or not the rare event has occurred. Local authorities who reply in the affirmative may be asked to supply the data. The sample will yield information concerning all other matters of local finance.

Since the sample will be designed on the basis of the T_h^2 values, it would be well to collect data from the entire population periodically, possibly in 5- or 10-year intervals. Local governments grow or decline; some taxes gain or lose importance; new taxes are levied; and new demands on receipts are made. All of these may call for new stratification or new allocation, or both.

Application of method illustrated

The following hypothetical case will illustrate the numerical calculations involved in the use of the preceding formulas. Assume that a State has 104 incorporated places of all sizes, that 7 of these have 10,000 or more people each, and that the population of the next largest community is about 7,000. It is desired to estimate total highway receipts and expenditures, as well as various subclasses of receipts and disbursements, for all incorporated places. The decision is made to include the seven larger places and to sample the smaller places. It is also decided that the objective upon which to base sample size should be an estimate of total capital outlay plus maintenance costs for highways in which there is confidence that in 2 chances out of 3 the sample estimate for the remaining 97 places will differ no more than 5 percent from the true value. In the preceding year, when no sample was drawn, this value was $1,205 (000). Assuming that this amount has not changed radically, the desired standard error of estimate, d, would equal (0.05) (1,205) or 60.25. The value of d^2 would be 3,630.0625. It is decided to stratify. Three strata are determined on the basis of the previous year's values of capital outlay plus maintenance expenditures. Table 4 gives the distribution of these values by strata.

Table 4.—Highway expenditures by incorporated places under 10,000 population in a hypothetical State

Highway expenditures for capital outlay and maintenance by incorporated places under 10,000 population		
Stratum 1, 63 incorporated places	Stratum 2, 24 incorporated places	Stratum 3, 10 incorporated places
$2,000	$10,500	$22,600
2,300	11,200	27,000
2,700	14,000	29,500
3,000	14,800	32,700
3,100	15,600	35,000
3,500	15,900	39,200
3,600	16,000	45,600
3,800	16,300	50,000
3,900	16,400	60,100
4,000	16,400	69,100
4,200	17,100	----------
4,300	17,200	----------
4,500	17,500	----------
4,600	17,700	----------
4,800	18,000	----------
4,800	18,100	----------
5,000	18,200	----------
5,000	18,300	----------
5,200	18,300	----------
5,200	18,500	----------
5,400	18,600	----------
5,500	18,600	----------
5,700	19,000	----------
5,800	19,500	----------
5,900	----------	----------
5,900	----------	----------
6,000	----------	----------
6,000	----------	----------
6,100	----------	----------
6,200	----------	----------
6,300	----------	----------
6,300	----------	----------
6,400	----------	----------
6,400	----------	----------
6,500	----------	----------
6,500	----------	----------
6,500	----------	----------
6,600	----------	----------
6,600	----------	----------
6,700	----------	----------
6,900	----------	----------
7,000	----------	----------
7,000	----------	----------
7,000	----------	----------
7,100	----------	----------
7,200	----------	----------
7,200	----------	----------
7,500	----------	----------
7,600	----------	----------
7,700	----------	----------
8,000	----------	----------
8,200	----------	----------
8,500	----------	----------
8,600	----------	----------
8,800	----------	----------
9,000	----------	----------
9,100	----------	----------
9,200	----------	----------
9,500	----------	----------
9,500	----------	----------
9,600	----------	----------
9,700	----------	----------
9,700	----------	----------
392,500 [1] (6,200)	401,700 [1] (16,700)	410,800 [1] (41,100)

[1] Average expenditures per incorporated place.

The formula and values for estimating the population variance of capital outlay plus maintenance expenditures for each stratum, based upon the previous year's values, are:

$$T_1^2=\frac{\sum_i^{63}\left(y_{1i}-\frac{\Sigma y_{1i}}{63}\right)^2}{62}$$

$$=\frac{63\sum^{63}(y_{1i})^2-\left(\sum^{63}y_{1i}\right)^2}{(63)(62)} \quad \ldots \text{(Stratum 1)}$$

$$=\frac{63\,(2{,}683.81)-(392.5)^2}{3{,}906}$$

$$=\frac{15{,}023.78}{3{,}906}=3.8463$$

$T_1=1.96$ (estimate of the standard deviation).

$$T_2^2=\frac{24\sum_i^{24}(y_{2i})^2-\left(\sum_i^{24}y_{2i}\right)^2}{(23)(24)}$$

$$=\frac{24(6{,}842.35)-(401.7)^2}{552} \quad \ldots \text{(Stratum 2)}$$

$$=\frac{2{,}853.51}{552}=5.1694$$

$T_2=2.27$

$$T_3^2=\frac{10\sum_i^{10}(y_{3i})^2-\left(\sum_i^{10}y_{3i}\right)^2}{(9)(10)}$$

$$=\frac{10(18{,}907.12)-(410.8)^2}{90} \quad \ldots \text{(Stratum 3)}$$

$$=\frac{20{,}314.56}{90}=225.7173$$

$T_3=15.02$

The following calculations lead to a choice between optimum allocation and proportional allocation:

$$\overline{T}=[1.96+2.27+15.02]\div 3=6.417-$$

$$(\overline{T})^2=41.18$$

$$\sigma^2_{T_h}=\frac{1}{97}\,[63(1.96-6.42)^2+24(2.27-6.42)^2+10(15.02-6.42)^2]$$

$$=\frac{1}{97}\,[2{,}406.1108]=24.8053$$

$$V^2_{T_h}=\frac{24.81}{41.18}=0.602$$

Therefore optimum allocation would be the preferred procedure.

A table similar to table 5 could be used to aid in the calculation of the sample size and in the allocation of the sample to the several strata.

General comments

The sample design was based upon one important characteristic of local finance, an estimate of total capital outlay plus maintenance costs for highways. Parenthetically, other characteristics that are highly correlated with the selected characteristic and which exceed it in dollar value, such as total highway expenditures, can be estimated from the sample with greater confidence than the selected design characteristic. Also, a different characteristic highly correlated with the design characteristic, if readily available, could have been used as the basis for stratification.

Seven incorporated places that could not conveniently be represented by other incorporated places were included with certainty in the total number of places to be selected. The remaining 97 incorporated places were grouped into three strata on the basis of the value of the design characteristic determined in the preceding year on a non-sample basis.

An estimate was calculated of the variance for each stratum on the basis of the previous year's value of the design characteristic. The square root of each of these values yielded an estimate of the standard deviation of the population of values in each stratum. The variance of these standard deviations about their average value was then calculated, and the ratio of this variance to the square of the average standard deviation was determined. Since this ratio was greater than one-third, optimum allocation was adopted rather than proportionate stratified sampling.

The calculations described in the preceding paragraph provided all values, except d^2, needed in the formulas for determining the total sample size and the allocation of this total among the three strata. The value of d^2 was fixed earlier when deciding on the sample design. It is the square of the range of values centered on the sample estimate which it is anticipated that the sample will yield. If the actual results do yield the design value of d, then the desired standard will have been attained.

Table 5.—Data for estimating sample size and strata allocations

Stratum	N_h	T_h^2	$N_h T_h^2$	T_h	$N_h T_h$	$\frac{(N_h T_h)}{328.16}$	n_h
1	63	3.8463	242.3169	1.96	123.48	0.376	7
2	24	5.1694	124.0656	2.27	54.48	.166	3
3	10	225.7173	2257.1730	15.02	150.20	.458	8
Total	97	------------	2623.5555	--------	328.16	1.000	[1] 18

[1] $n=(328.16)^2/[3630.0625+2623.5555]=17.2$ (rounded to 18).

Motor-Vehicle Size and Weight Limits

A comparison of State legal limits of motor-vehicle sizes and weights with standards recommended by the American Association of State Highway Officials is given in a table on pages 250–251. The statutory limits reported in this tabulation, prepared by the Bureau of Public Roads as of July 1, 1959, have been reviewed for accuracy by the appropriate State officials.

Statutory limits are shown for width, height, and length of vehicles; number of towed units; maximum axle loads for single and tandem axles; and maximum gross weights for single-unit trucks, truck-tractor semitrailer combinations, and other combinations.

Factors Influencing Mass-Transit and Automobile Travel in Urban Areas

BY THE DIVISION OF HIGHWAY PLANNING
BUREAU OF PUBLIC ROADS

Reported by WARREN T. ADAMS, Highway Transport Research Engineer

The objective of the research discussed in this article is to develop a relationship between the use of public and private transportation in urban areas and the principal factors influencing that use. If that can be satisfactorily done, an estimate can be made of the usage of each form of transportation under a given set of influencing factors and, as a consequence, estimates can be made of the total construction and operating costs involved in providing the required transportation system, including terminal facilities for each travel mode. The final step is to estimate the benefits of each plan or program in relation to cost.

By means of multiple regression analysis, data for 16 cities were used to develop a relationship between relative urban travel mode use and its principal influencing factors. Subsequently, this relationship was successfully tested through application of the derived equation to data received from five additional cities. The establishment of this relationship should materially aid in urban transportation planning, as the use of the equation makes it possible to predict relative transit use with an acceptable degree of accuracy.

A similar relationship between relative mode use and the principal influencing factors is also being investigated for major subdivisions of each urbanized area. Preliminary analysis indicates that the relationship for the subdivisions will be similar to that for the whole urbanized area.

WITHIN the next 25 years, there undoubtedly will be changes in modes of urban transportation. It is quite doubtful, however, that drastic changes will take place so rapidly that the basis for planning will be nullified overnight. At least so far as can be foreseen, the establishment of the relationship between the use of public and private transportation will permit the preparation of transportation plans on a more realistic basis.

A start has been made in establishing this needed relationship. The results of the investigation described here indicate that the approach will yield relationships, sufficiently accurate for transportation planning, that will enable planning officials to predict the travel mode split for a specific set of conditions.

In most urban areas since the late 1920's and excluding the depression and war years, the automobile has been supplanting public transportation as the mass carrier of people. The rate of change, however, has not been uniform from city to city or from year to year. In some metropolitan areas, principally the larger cities east of the Mississippi River, the decline in transit riding apparently is tending to level off. The trend from transit to automobile, however, has been continuing at either an accelerated or constant rate in many cities across the nation.

Public and private organizations have long been seeking a means for determining the change in both relative use and actual use of transit and automobile in urban areas. Transit and regulatory agency officials need this information to establish an equitable fare structure and to provide adequate transit service. Highway engineers have been faced with the necessity of determining whether the cost of street improvements, parking facilities, and transit facilities on highways will be justified by future use. Planning and municipal agencies are concerned with the interrelated effect of land-use distribution and population on the one hand, and the use of transit and automobile with the accompanying effect on public utilities on the other. Municipal authorities must also determine the impact on the tax structure of relative travel mode use, especially if transit operations should become the responsibility of a public agency. Primarily, all of these groups are concerned with the factors and elements that cause variations in mode of travel and their relation to the economic and political system.

Study Procedure

In previous studies, attempts were made to establish relationships between transit-riding habits and several related factors. In general, these factors were population, automobile registrations, transit service, economic welfare, and transportation costs. The studies, however, have not yielded conclusive results. They did not develop relationships that would make it possible to forecast relative mode use with an acceptable degree of accuracy.

In the last few years, more attention has been given to land use as one of the principal factors affecting urban transportation. Studies have indicated an appreciable degree of correlation between travel mode on the one hand, and residential, commercial, and industrial land use on the other.

In this study, a land-use distribution factor has been combined with factors relating to population, automobile ownership, employment, dwelling units, transit-service ratio and size of urbanized area in order to develop a basis for forecasting travel mode use.

Source of Data

The home-interview origin-and-destination survey reports for all cities completed since January 1948 were reviewed; cities not having transit service were naturally eliminated. Cities in which surveys were made prior to 1948 were not considered, principally because war-caused distortions still exerted a major influence through 1947, and the quality of many of the early surveys was questionable. Of the 30 cities studied, 8 surveys were completed in 1948, 7 in 1949, 6 in 1950, 3 in 1951–52, and 6 in 1953.

Through the cooperation of the American Transit Association, information on transit service at the time of the origin and destination survey was obtained from the transit companies in 22 of the 30 cities. Since previous studies had indicated that one of the key factors in any developed relationship might be land-use distribution, such information corresponding to the origin-and-destination survey period was also requested from the planning agencies of each city.

Methods of Analysis

Simple and multiple regression equations were developed and tested in the study of the relationship of travel mode to various factors. Among the simple factors examined were population, automobile ownership, trips to work, and total survey land area. Among the compound factors investigated were population density, automobile ownership per capita, and employment per capita. Also tested were the regression relationships between travel mode split and factor combinations, including various compound factors such as employment and automobile ownership per capita. None of these relationships yielded either an acceptable standard error of estimate or a high degree of correlation.

Symbols and definitions

Letter symbols used to represent the various factors discussed in subsequent sections of this article are defined as follows:

A_c = Commercial and industrial land area (sq. mi.) within a 1-mile radius of the CBD center.

A_{tc} = Commercial and industrial land area (sq. mi.) within the entire urbanized area.

a = Number of automobiles owned.

D = Parking demand as related to the number of types of employment.

E = Economic factor as determined by $(P/e)^{3.5}(h/e)^{1.5}(P/a)^{1.0}(h/a)^{1.5}$.

e = Number of employees going to work on an average weekday.

F = Terminal or parking facility factor.

h = Number of dwelling units.

M = Urbanized land area in square miles.

P = Population over 5 years of age in the survey urbanized area.

R_{1c} = Mean distance of commercial and industrial land within a 1-mile radius of the CBD center.

R_{tc} = Mean distance from the CBD center of commercial and industrial land within the entire urbanized area.

R_p = Mean distance of center of population from the CBD center.

R_u = Mean distance of entire urbanized area from the CBD center.

r_1 = Land-use distribution ratio, $1-(R_p/R_u)$.

r_2 = Land-use distribution ratio, R_pR_{tc}/R_u.

r_3 = Land-use distribution ratio, A_{1c}/A_{tc}.

r_4 = Land-use distribution ratio, $(A_{tc}/A_{tc}R_{tc}) = (1/R_{tc})$.

r_5 = Land-use distribution ratio, $(A_{1c}R_{1c}/A_{tc}R_{tc})$.

S_r = Ratio of the square root of the average vehicle speed for the different travel modes.

T = Transit-service ratio factor as determined by $[V^{1.0}/P^{1.50}M^{0.25}](S_r)(F/D)$.

U = Land-use distribution factor as determined by $(r_1)(r_4)(r_5)$.

V = Equivalent revenue vehicle-miles operated per weekday during the survey.

y = Reported percentage of total person trips made via transit.

y_1 = Estimated percentage of total person trips made via transit.

Development of equations

As transit service information was received, tests were made to determine whether there might be a significant relationship between transit service data and travel mode split. As the lone independent variable, transit service did not produce a satisfactory estimating equation. Using a transit-service ratio factor in conjunction with a population factor, a combination automobile and employment factor, and an urbanized land-use factor, semilog multiple variable equations of the form $y_1 = A + b_1 \log P + b_2 \log E + b_3 \log T + b_4 \log M$ were developed that gave the results shown in figure 1 for the cities with available data for 1948, 1949, and 1950.

These results indicated that there might be at least one other factor which, if included, would produce the relationship sought regardless of the year. From previous studies, it was believed that this factor might be largely based on land-use distribution, for with each succeeding year since 1948 there has been an increasing decentralization of residential, commercial, and industrial land use with respect to the central business district (CBD). Tests were made of ranges of various ratios and combinations of ratios involving distribution of land used commercially and industrially within and about the CBD, population distribution with respect to the center of the urbanized land area, and population distribution with respect to employment location. Although there were differences in the amount of variation explained, nearly all combinations tended to reduce the year-to-year variations. Using the semilog equation $y_1 = A + b_1 \log P + b_2 \log E + b_3 \log T + b_4 \log U + b_5 \log M$, the land-use distribution factor, U, brought the 16 cities listed in table 1 into a straight-line relationship as shown in figure 2.

As additional information is obtained, the land-use distribution factor as well as other factors in use may require modifications.

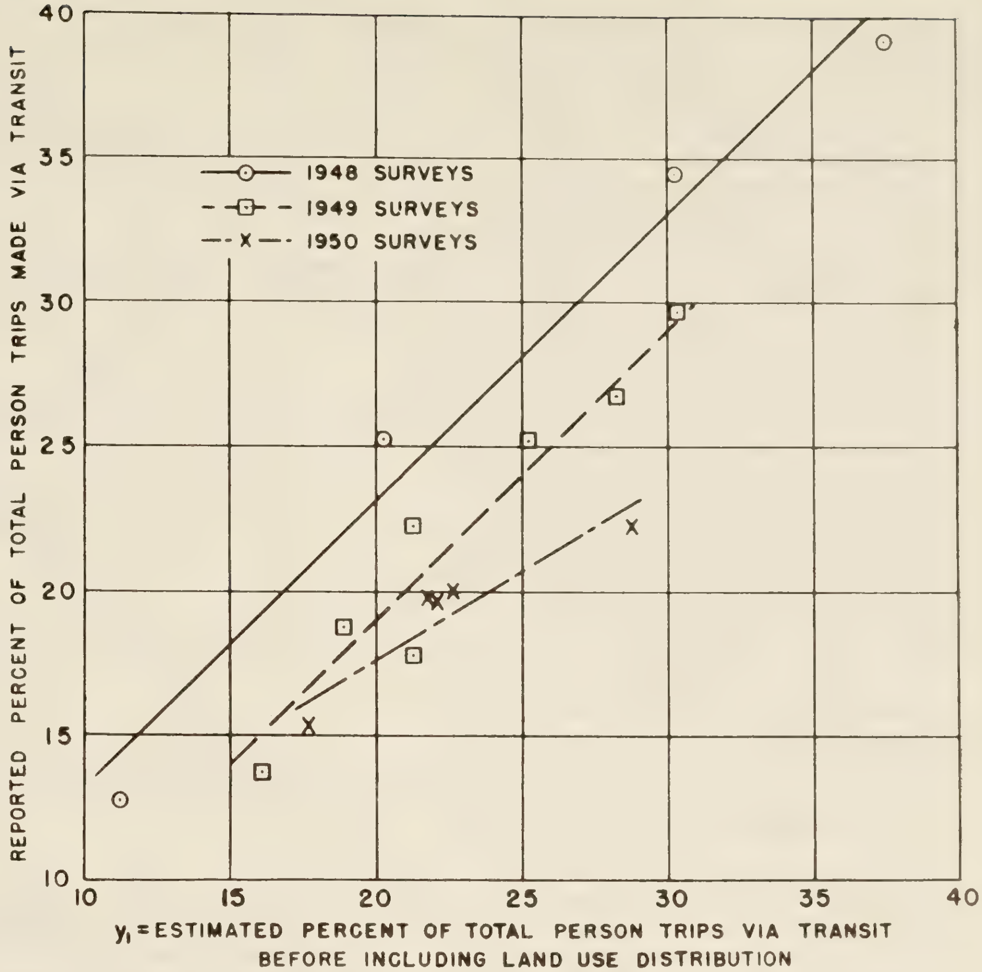


Figure 1.—Relationship of reported relative transit use to that derived from estimating equation, before inclusion of the land-use distribution factor.

Table 1.—Basic data for 16 cities used in developing relative transit use equation for entire urbanized area

O-D survey year and urbanized area	Population over 5 years of age *P*	Dwelling units *h*	Employees going to work per average weekday *e*	Automobiles owned *a*	Equivalent revenue vehicle-miles per average weekday *V*	Urbanized land area *M*	Land-use distribution factor *U*	Reported relative transit use *y*
	Thousands	*Thousands*	*Thousands*	*Thousands*	*Miles*	*Sq. mi.*	*Ratio*	*Percent*
1948:								
Washington, D.C	992.6	336.2	380.8	203.5	155,060	108.80	0.00456	39.3
Allentown-Bethlehem, Pa	156.4	48.8	61.6	28.6	12,490	18.95	.00628	32.1
Tacoma, Wash	125.0	48.0	36.1	35.2	11,960	45.80	.03350	25.2
Tucson, Ariz	113.7	38.7	29.6	32.9	3,610	41.00	.07398	12.7
1949:								
Albuquerque, N. Mex	100.8	34.9	30.4	27.5	4,510	38.93	.02035	13.6
Madison, Wis	94.3	33.4	29.6	25.3	5,780	20.40	.03034	22.4
Racine, Wis	69.5	23.3	23.1	18.5	3,210	11.56	.06204	17.8
Sharon-Farrell, Pa	44.3	13.7	12.9	9.4	1,530	8.08	.19520	18.5
1950:								
Dallas, Tex	471.1	168.1	182.5	153.8	57,900	173.20	.00253	19.8
Rockford, Ill	102.5	36.2	48.7	33.1	6,020	24.54	.05248	15.1
Altoona, Pa	77.5	24.1	17.1	16.8	3,680	15.82	.03204	20.0
Kenosha, Wis	50.2	17.3	21.7	13.4	2,580	8.32	.19264	19.6
1953:								
Detroit, Mich	2,642.2	875.4	971.1	845.8	183,170	548.85	.00004	16.7
Houston, Tex	765.9	272.7	278.8	256.3	48,020	244.50	.00024	13.6
San Diego, Calif	513.0	175.0	152.1	250.5	36,170	152.01	.00025	13.2
Stockton, Calif	124.3	41.6	34.9	70.0	3,000	31.89	.01109	7.8

Variations in some factors have been investigated, and several satisfactory estimating equations with only minor variations have been developed. All equations have yielded a standard error of estimate of less than 1.5 percentage points, and several have yielded less than 1.0 percentage point of the reported transit use as a percentage of total person trips for the entire urbanized areas of the 16 cities. Transit usage as reported in table 1 for these cities ranged from 8 to 40 percent of the total trips made by persons, with a mean of approximately 20 percent. Thus, a standard error of 1.5 percentage points is equivalent to 7.5 percent of the mean value of 20 percent for total transit trips per weekday.

Of the semilog, multiple regression equations developed, the following equation was used in this study: $y_1 = -2.6466 + 3.7084 \log P + 0.3912 \log E + 2.3757 \log T + 0.4918 \log U - 0.9708 \log M$. The basic data for this equation are found in table 1.

Since developing the estimating equation, complete information has been obtained from 5 cities in addition to the 16 cities listed in table 1. Applying the equation to the data for the five cities, the results, as illustrated in figure 3, were found to be within the previously stated standard error of estimate.

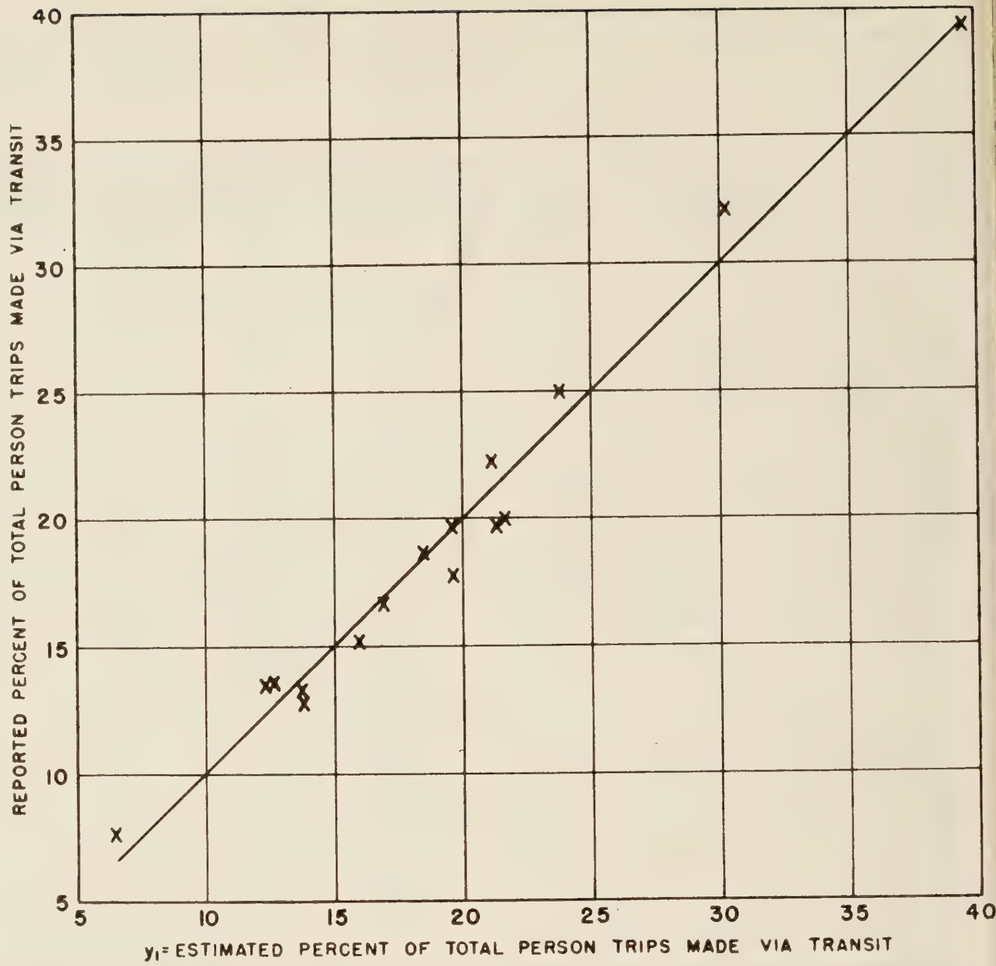


Figure 2.—Relationship of reported relative transit use to that derived from the estimating equation, after inclusion of the land-use distribution factor.

Discussion of Factors

In the estimating equation, the three compound factors—economic, transit service, and land use—have been developed through testing variations of each over a range that would determine the maximum effect of each variable in correlation with other potential variables. As more information is obtained from the cities mentioned in this article as well as additional cities, and as this information is adapted to electronic computer programing, it is anticipated that more precise parameters will be established.

Economic factor, *E*

Apparently there is a high degree of correlation between relative use of each transportation mode and some economic factor. Many contend that this factor is either income or wealth. But what income or wealth? Is it gross or net? What should be included and what deducted? Moreover, how could accurate measurements of income or wealth be made? Correlating travel mode use and related origin and destination information with sufficiently accurate income or wealth data will be most difficult under legal restrictions applying to the release of the latter.

There may be other economic items that have a higher degree of correlation with travel mode than the ones used in this study—population, number of dwelling units, number of automobiles owned, and number of employees going to work on an average weekday. These four items seem to be the best available that can be accurately measured. The simple linear correlation coefficient relating the various compounds of the components of E (economic factor) with relative mode use varied between 0.40 and 0.60.

The use of both population and number of dwelling units in relation to the number of automobiles owned and the number of employees going to work per average weekday may be challenged. The correlation obtained by use of these factors in combination has been greater than when only one factor has been used. This may be due to compensating errors in the origin and destination surveys, and to the effect of differences in population per dwelling unit.

Transit-service ratio factor, *T*

Although not approaching unity, there is a significant degree of correlation between the developed transit service variable and the dependent transit use variable. The simple linear correlation coefficient for the transit-service ratio factor and relative transit use has ranged from 0.30 to 0.45. The degree of correlation not only varies from city to city, but also within many cities there is likely to be an even greater variation among subdivisions (sectors, districts, etc.) or transportation channels. Furthermore, the effect of each of the components in the transit service factor apparently varies among cities and their subdivisions.

Equivalent revenue vehicle-miles, V.—Equivalent revenue vehicle-miles operated per weekday are expressed in terms of a 50-seat bus revenue-mile. Included are all revenue vehicle-miles operated per weekday, regardless of the number of passengers carried on each vehicle trip. This item has been derived by applying a carrying capacity factor to the average weekday revenue vehicle-miles recorded during the survey. The capacity factor has been developed through assignment of each vehicle size, by time periods, in proportion to the ages of the active vehicle groups. Inasmuch as it is impossible to obtain actual average carrying capacity during the surveys without a prior uniform arrangement with the transit operators, this derivation gives an arbitrary, but uniform, estimate for all cities that most nearly approaches the "true" average.

Average vehicle-speed ratio, S_r.—The ratio of the square root of the speeds of automobile and transit vehicle travel is one of the components of the transit-service ratio factor. There are few who do not consider this ratio an influencing item in determining the travel mode split. Yet based on available data, the standard error is increased by only 0.1 percent when the speed ratio factor is excluded. This would indicate that the variation of this ratio from city to city is not appreciable with respect to the entire urbanized area for each city. Due to the limited amount of data available and the apparent relatively small effect on the standard error, the speed ratio factor has not been included in the estimating equation adopted for this article. The investigation so far has borne out that the use of the ratio of the square root of the speed of each travel mode, although

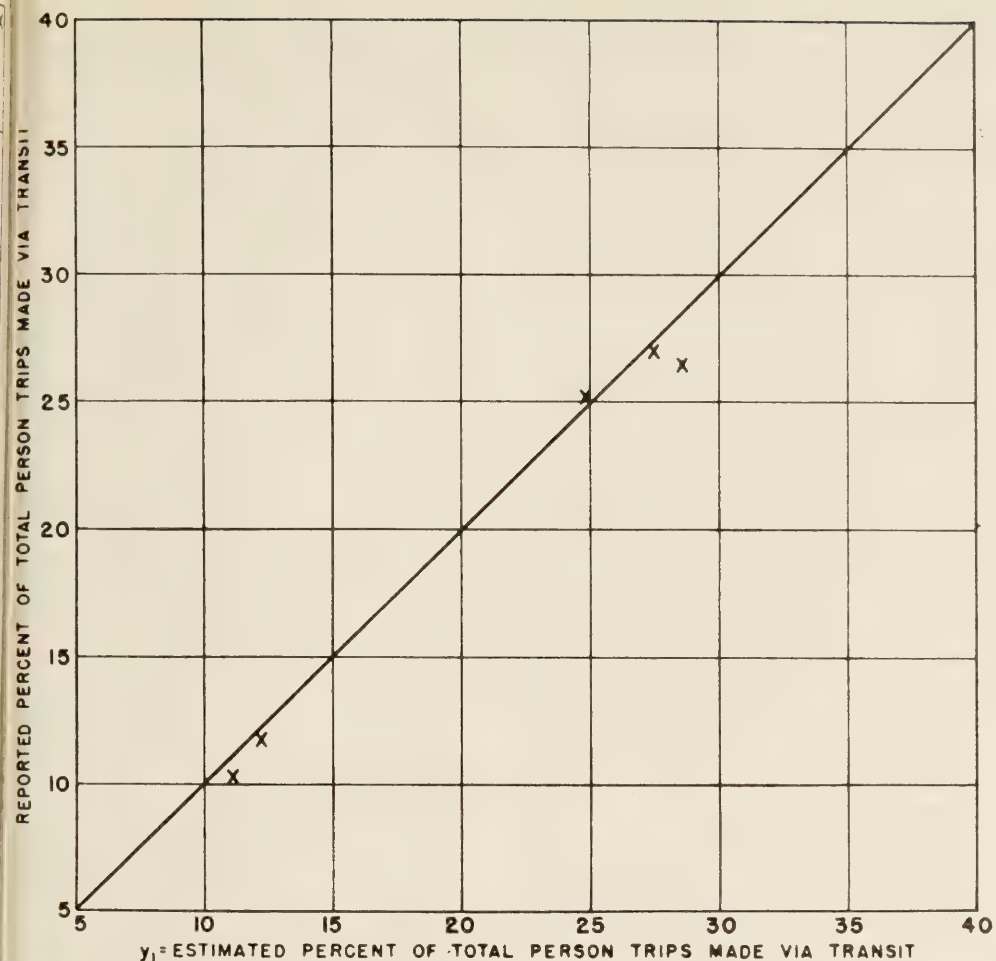


Figure 3.—Relationship of reported relative transit use to estimated relative use in five cities.

reducing the sensitivity of the component, appears to be most satisfactory.

The limited work done to date by subdivision of urban areas has pointed to a much greater influence of the speed component for subdivisions and transportation channels than for entire urban areas. This is accounted for by the greater spread of relative speeds within the subdivisions. The results still indicate, however, that the ratio should be based on the square root of the respective speeds.

It is quite possible that additional data may establish that different ratios should be applied to the two principal components of overall travel time for each travel mode, namely, vehicle speed and the terminal factor. The effect of these components must be more accurately determined not only to estimate transit and automobile usage under specific conditions, but also to develop the required transit, parking, and highway capacities with attendant capital and operating costs for the estimated use of each travel mode.

Land-use distribution factor, *U*

The land-use distribution factor is a complex one that has been developed from a series of studies within the limitations of available material, time, and computing equipment. In relating this factor to travel mode use, the simple linear correlation coefficient varied from 0.60 to 0.75. It appears likely that more efficient analysis of present and future data by means of an electronic computer program will produce either more precise values for the factors now being used or simpler factors that may prove to be more satisfactory. For the entire urban area and for the subdivisions investigated, there appear to be five land-use distribution ratios about the CBD center that should be considered. These ratios are listed on page 257 as r_1, r_2, r_3, r_4, and r_5.

In arriving at the mean distance to the CBD center from the centroid of each of the items included in the land-use ratios, the same procedure was used for each item. Therefore, a detailed description of the derivation of one item (commercial and industrial land use) will suffice for all. Each city was divided into four quadrants by rectangular coordinate axes passing through the CBD center. In each quadrant, the area of each industrial and commercial parcel, or each group of adjacent parcels, actually used for one of these purposes at the time of the origin and destination survey was multiplied by the distance from the CBD center to the centroid of the parcel or group of parcels. These products were then summed for the four quadrants, and this summation was divided by the summation of the areas of all the industrial and commercial parcels in the urbanized area.

In some instances it was found to be more efficient to determine the mean distance through summing the products obtained by multiplying the areas by their distances to the two coordinate axes, and then extracting the square root of the sum of the squares of those two product summations.

In the estimating equation adopted, ratios r_2 and r_3 were not included in the land-use distribution factor. Studies which have not been concluded indicate, however, that the inclusion of ratios r_2 and r_3 with possible modification of the other ratios would reduce the standard error of estimate.

The present study has shown that it is not necessary to differentiate between commercial and industrial land for the entire urban area. This is apparently due to the balancing effect of the two land uses over a complete urban area. Within highly specialized subdivisions or transportation channels serving predominantly one type of land use, the investigation shows that the two uses will have to be considered separately. It may even be necessary to subdivide the two classifications into four classes: industrial, office, shopping durable, and shopping service and convenience. Based on probable accuracy of land-use forecasts, it is questionable whether further breakdowns can be justified for transportation channel subdivisions and even much less justifiable for subdivisions comparable to origin and destination districts. To justify more classifications of land use, a much greater specialization of land use in a transportation channel subdivision would be required than has been found or seems probable in the future.

Urbanized land area factor, *M*

Defining urbanized area boundaries is of utmost importance in determining the relative use of transit vehicles and automobiles. Urbanized areas, as used in the present study, have been confined to contiguously developed land, and future estimates must be based on anticipated contiguously developed land. Furthermore, such land must have a minimum residential population per area unit of 500 persons per square mile, or a minimum number of 2,000 trip-ends per square mile to be included as part of the urbanized area. Islands of vacant land should be included if the land outside is sufficiently developed to bring the combination of vacant land and adjacent, outside developed land up to the minimum requirements just stated. Pockets of vacant land at the origin and destination boundary not meeting these specifications should be excluded.

Many subdivisions with population or trip-ends above the minimum requirements cannot be served by transit unless the service costs are partially defrayed by either the subdivision, the entire urban area, or the entire transit system. In border subareas where either the residential population or number of trip-ends is less than the minimum, the only mode of urban transportation will be the privately owned automobile unless transit service is furnished by an intercity carrier, or it is almost entirely underwritten on a service charge basis due to relatively low transit use. If border land with subminimum population or trip-ends is included in the study, the urbanized area factor, M, would be appreciably changed.

Work in Progress

Of the factors investigated so far, the three that contribute most in explaining the variance are the transit-service ratio factor, the land-use distribution factor, and the economic factor. Some of the factors are still being studied to determine whether they should be modified or replaced. It is possible that the estimating equation may be appreciably changed as a result of these continuing studies. If the work in the past, however, is a criterion for the future, the estimating equation will not likely be altered significantly.

Expansion of project

In addition to continuing research on the whole metropolitan area, this study is now being expanded in two directions. Generalizing, it can be stated that the present estimating equation is an expression of the division of trips between transit and automobile in relation to factors pertaining to the home area (employment, automobile ownership, and population distribution), and to factors applying to the entire metropolitan area (land-use distribution, transit-service ratio, total population, and urbanized land area). The equation is presently being tested to determine whether it, or a modification thereof, will apply to subdivisions of each metropolitan area. At the present time, sufficient information to investigate such application has been obtained from only three cities.

The results so far indicate that the equation, after modifying the land-use and transit-service ratio factors, will make it possible to forecast with acceptable accuracy the split between transit and automobile trips for each subdivision. Two items that apparently have more influence on the travel mode within each subdivision than for the whole urban area are the average ratio of overall trip time by the two transportation modes, and the ratio of commercial to industrial land use. Since only 20 subdivisions have been partially investigated, a precise basis for modifying the transit-service ratio and land-use factors has not yet been developed.

The other extension of the project has been an attempt to determine the influence on the estimating equation of other factors in destination areas. In this subphase, even less information is available. A limited amount of information on CBD destination factors has been gathered in several cities, and the relation between these factors and the travel mode split has been tested. Foremost among the items that apparently should be introduced into the equation is a parking facility factor. The equation, modified by this factor, appears to yield a low standard error of estimate in predicting the travel mode split in destination areas. The factor, however, is not confined to total parking space in each destination area; it also includes accessibility to demand as expressed by a relationship including parking charges and walking and parking time.

Areas of future research

Only the surface has been scratched in attempting to establish factors and estimating equations pertaining to the travel mode split by home and destination zones. Relationships by subdivision should also be developed for the peak travel period. Some work is under way on this peak period relationship for the entire metropolitan area.

The results of research accomplished so far seem to indicate that the speed factor, in general, varies as the ratio of the square root of trip speed. Convenience and irritation items, to the extent that it has been possible to measure them, and modified somewhat by cost, are apparently as important, if not more so, than absolute vehicle speed. This observation, however, may not hold for freeway and rapid-transit operation. In fact, testing additional data for vehicle operation on unrestricted rights-of-way may alter the findings in this field. The analysis of the Chicago origin and destination survey should yield much information on this phase. It is the only city with both rapid transit and a limited amount of freeway traffic data now available for testing.

To carry out further analyses of relationships, much additional information will be needed. Many cities and transit companies have cooperated. If the needed relationships are to be established, it will be necessary to call on these and other cities for more basic data from time to time. Much of the information should be gathered at the time of the origin and destination survey; in fact, it should be made a part of it. It will take time and money to gather and assemble the information, but it should be beneficial to both city planning agencies and transit companies.

Selected Bibliography

Bacon, Edmund N. *Urban Redevelopment and Highway Planning.* Highway Research Board Bulletin 64, Washington, 1952. pp. 9–12.

Detroit Metropolitan Area Traffic Stud Part I, Data Summary and Interpretatio J. D. Carroll, Jr., study director. Michiga State Highway Department, Lansing, Jul 1955. 162 pp.

Detroit Metropolitan Area Traffic Stud Part II, Future Traffic and a Long Ran Expressway Plan. J. D. Carroll, Jr., stud director. Michigan State Highway De partment, Lansing, Mar. 1956. 157 pp.

Hoyt, Homer. *Changing Land-Use Pattern as a Basis for Long-Range Highway Plan ning.* Highway Research Board Bulleti 64, Washington, 1952. pp. 1–8.

Mertz, William L., and Hamner, Lamelle B. *A Study of Factors Related to Urban Trave* PUBLIC ROADS, vol. 29, No. 7, Apr. 1957 pp. 170–174.

Relationships of Measures of Intra-Urba Accessibility to Intensity and Type of Lan Use. RESEARCH DIGEST. University o Illinois, Department of City Planning an Landscape Architecture, Champaign, vol. 4 No. 1, May 1957.

This is a research project undertake by the staff of Division VII of the Chicag Area Transportation Study. Findings o this study will be incorporated in the pub lished report (in process) of the Chicag Area Transportation Study. Interim prog ress reports are included in C.A.T.S RESEARCH NEWS, a fortnightly publicatio of the Chicago Area Transportation Study

Sharpe, Gordon B., Hansen, Walter G., an Hamner, Lamelle B. *Factors Affectin Trip Generation of Residential Land-Us Areas.* PUBLIC ROADS, vol. 30, No. 4 Oct. 1958. pp. 88–99.

Stephenson, Charles A. *Balanced Circulation for Metropolitan Areas.* MASS TRANSPORTATION, vol. 51, No. 11, Nov. 1955. pp. 19–25.

———. *Passenger Seat Miles Versus Vehicle Miles.* MASS TRANSPORTATION, vol. 42, No. 12, Dec. 1946. pp. 535–536.

———. *Transit's Prospects for Postwar Traffic.* American Transit Association Convention-in-Print, 2d sess., Nov. 3, 1944. pp. 10–12.

Transit Riding, Revenues and Fare Structures. Final report on research project No. 4. American Transit Association, New York, Aug. 1949, 161 pp.

Wynn, F. Houston. *Intracity Traffic Movements.* Highway Research Board Bulletin 119, Washington, 1955. pp. 53–68.

———. *Is This Trip Necessary?* PLANNING 1955. American Society of Planning Officials, Chicago, Apr. 1956. pp. 140–143.

U. S. GOVERNMENT PRINTING OFFICE: 1959

A list of the more important articles in PUBLIC ROADS may be obtained upon request addressed to Bureau of Public Roads, Washington 25, D.C.

PUBLICATIONS of the Bureau of Public Roads

The following publications are sold by the Superintendent of Documents, Government Printing Office, Washington 25, D.C. Orders should be sent direct to the Superintendent of Documents. Prepayment is required.

ANNUAL REPORTS

Annual Reports of the Bureau of Public Roads:

1950 (out of print).	1955, 25 cents.
1951, 35 cents.	1956 (out of print).
1952, 25 cents.	1957 (out of print).
1953 (out of print).	1958, 30 cents.
1954 (out of print).	

REPORTS TO CONGRESS

A Report of Factors for Use in Apportioning Funds for the National System of Interstate and Defense Highways, House Document No. 300 (1958). 15 cents.

Consideration for Reimbursement for Certain Highways on the Interstate System, House Document No. 301 (1958). 15 cents.

Factual Discussion of Motortruck Operation, Regulation, and Taxation (1951). 30 cents.

Federal Role in Highway Safety, House Document No. 93 (1959). 60 cents.

First Progress Report of the Highway Cost Allocation Study, House Document No. 106 (1957). 35 cents.

Highway Needs of the National Defense, House Document No. 249 (1949). 50 cents.

Interregional Highways, House Document No. 379 (1944). 75 cents.

Local Rural Road Problem (1950). 20 cents.

Needs of the Highway Systems, 1955–84, House Document No. 120 (1955). 15 cents.

Progress and Feasibility of Toll Roads and Their Relation to the Federal-Aid Program, House Document No. 139 (1955). 15 cents.

Progress Report on the Federal-Aid Highway Program, House Document No. 74 (1959). 70 cents.

Public Utility Relocation Incident to Highway Improvement, House Document No. 127 (1955). 25 cents.

Third Progress Report of the Highway Cost Allocation Study, House Document No. 91 (1959). 35 cents.

PUBLICATIONS

Bibliography of Highway Planning Reports (1950). 30 cents.

Braking Performance of Motor Vehicles (1954). Out of print.

Catalog of Highway Bridge Plans (1959). $1.00.

Construction of Private Driveways, No. 272MP (1937). 15 cents.

Criteria for Prestressed Concrete Bridges (1954). 15 cents.

Design Capacity Charts for Signalized Street and Highway Intersections (reprint from PUBLIC ROADS, Feb. 1951). 25 cents.

PUBLICATIONS (Continued)

Financing of Highways by Counties and Local Rural Governments, 1942–51. 75 cents.

General Location of the National System of Interstate Highways, Including All Additional Routes at Urban Areas Designated in September 1955. 55 cents.

Highway Bond Calculations (1936). 10 cents.

Highway Capacity Manual (1950). $1.00.

Highway Statistics (published annually since 1945):
- 1955, $1.00.
- 1956, $1.00.
- 1957, $1.25.

Highway Statistics, Summary to 1955. $1.00.

Highway of History (1939). 25 cents.

Identification of Rock Types (reprint from PUBLIC ROADS, June 1950). Out of print.

Legal Aspects of Controlling Highway Access (1945). 15 cents.

Manual on Uniform Traffic Control Devices for Streets and Highways (1948) (*including 1954 revisions supplement*). $1.25.

Revisions to the Manual on Uniform Traffic Control Devices for Streets and Highways (1954). *Separate,* 15 cents.

Parking Guide for Cities (1956). 55 cents.

Public Control of Highway Access and Roadside Development (1947). 35 cents.

Public Land Acquisition for Highway Purposes (1943). 10 cents.

Results of Physical Tests of Road-Building Aggregate (1953). $1.00.

Selected Bibliography on Highway Finance (1951). 60 cents.

Specifications for Aerial Surveys and Mapping by Photogrammetric Methods for Highways, 1958: a reference guide outline. 75 cents.

Standard Specifications for Construction of Roads and Bridges on Federal Highway Projects, FP–57 (1957). $2.00.

Standard Plans for Highway Bridge Superstructures (1956). $1.75.

Transition Curves for Highways (1940). $1.75.

Single copies of the following publications are available upon request addressed to the Bureau of Public Roads. They are not sold by the Superintendent of Documents.

Indexes to PUBLIC ROADS, volumes 17–19 and 23.

Title Sheets for PUBLIC ROADS, volumes 24–29.

UNITED STATES
GOVERNMENT PRINTING OFFICE
DIVISION OF PUBLIC DOCUMENTS
WASHINGTON 25, D.C.

OFFICIAL BUSINESS

PENALTY FOR PRIVATE USE TO AVOID PAYMENT OF POSTAGE, $300
(GPO)

If you do not desire to continue to receive this publication, please CHECK HERE ☐; tear off this label and return it to the above address. Your name will then be removed promptly from the appropriate mailing list.