2003 SURVEY RESULTS

APPENDIX A: *PRECISION OF SAMPLE ESTIMATES

***Reprinted from:**

Boyle, J. and P. Vanderwolf. 2003 Motor Vehicle Occupant Safety Survey. Volume I. Methodology Report. Washington, DC: U.S. Department of Transportation, National Highway Traffic Safety Administration.

Precision of Sample Estimates

The objective of the sampling procedures used on this study was to produce a random sample of the target population. A random sample shares the same properties and characteristics of the total population from which it is drawn, subject to a certain level of sampling error. This means that with a properly drawn sample we can make statements about the properties and characteristics of the total population within certain specified limits of certainty and sampling variability.

The confidence interval for sample estimates of population proportions, using simple random sampling without replacement, is calculated by the following formula:

$$z * \left[se(x) = \sqrt{\frac{(p * q)}{(n-1)}} \right]$$

Where:

se (x)	=	the standard error of the sample estimate for a proportion;
p	=	some proportion of the sample displaying a certain characteristic or attribute;
q	=	(1 - p);
n	=	the size of the sample;
z	=	the standardized normal variable, given a specified confidence level (1.96 for samples of this size).

The sample sizes for the surveys are large enough to permit estimates for sub-samples of particular interest. Table 56, on the next page, presents the expected size of the sampling error for specified sample sizes of 12,000 and less, at different response distributions on a categorical variable. As the table shows, larger samples produce smaller expected sampling variances, but there is a constantly declining marginal utility of variance reduction per sample size increase.

TABLE 56
Expected Sampling Error (Plus Or Minus)
At The 95% Confidence Level
(Simple Random Sample)

Size of Sample or Subsample	Percentage of the Sample or Subsample Giving A Certain Response or Displaying a Certain Characteristic for Percentages Near: 10 or 90	20 or 80	30 or 70	40 or 60	50
12,000	0.5	0.7	0.8	0.9	0.9
8,000	0.7	0.9	1.0	1.1	1.1
6,000	0.8	1.0	1.2	1.2	1.3
4,500	0.9	1.2	1.3	1.4	1.5
4,000	0.9	1.2	1.4	1.5	1.5
3,000	1.1	1.4	1.6	1.8	1.8
2,000	1.3	1.8	2.0	2.1	2.2
1,500	1.5	2.0	2.3	2.5	2.5
1,300	1.6	2.2	2.5	2.7	2.7
1,200	1.7	2.3	2.6	2.8	2.8
1,100	1.8	2.4	2.7	2.9	3.0
1,000	1.9	2.5	2.8	3.0	3.1
900	2.0	2.6	3.0	3.2	3.3
800	2.1	2.8	3.2	3.4	3.5
700	2.2	3.0	3.4	3.6	3.7
600	2.4	3.2	3.7	3.9	4.0
500	2.6	3.5	4.0	4.3	4.4
400	2.9	3.9	4.5	4.8	4.9
300	3.4	4.5	5.2	5.6	5.7
200	4.2	5.6	6.4	6.8	6.9
150	4.8	6.4	7.4	7.9	8.0
100	5.9	7.9	9.0	9.7	9.8
75	6.8	9.1	10.4	11.2	11.4
50	8.4	11.2	12.8	13.7	14.0

NOTE: Entries are expressed as percentage points (+ or -)

However, the sampling design for this study included a separate, concurrently administered over-sample of youth and young adults (age 16-39). Both the cross-sectional sample and the over-sample of the youth/younger adult population were drawn as simple random samples; however, the disproportionate sampling of the age 16-39 population introduces a design effect that makes it inappropriate to assume that the sampling error for total sample estimates will be identical to those of a simple random sample.

In order to calculate a specific interval for estimates from a sample, the appropriate statistical formula for calculating the allowance for sampling error (at a 95% confidence interval) in a stratified sample with a disproportionate design is:

$$\mathrm{ASE} = 1.96\sqrt{\sum_{h=1}^{g}\left[\mathrm{W}_h^2\left\{(1-f_h)\left(\frac{s_h^2}{n_h-1}\right)\right\}\right]}$$

where:

ASE	=	allowance for sampling error at the 95% confidence level;
h	=	a sample stratum;
g	=	number of sample strata;
W_h	=	stratum h as a proportion of total population;
f_h	=	the sampling fraction for group h - the number in the sample divided by the number in the universe;
s^2_h	=	the variance in the stratum h - for proportions this is equal to p_h (1.0 - p_h);
n_h	=	the sample size for the stratum h.

Although Table 56 provides a useful approximation of the magnitude of expected sampling error, precise calculation of allowances for sampling error requires the use of this formula. To assess the design effect for sample estimates, we calculated sampling errors for the disproportionate sample for a number of key variables using the above formula. These estimates were then compared to the sampling errors for the same variables, assuming a simple random sample of the same size. The two strata (h^1 and h^2) in the disproportionate sample were all respondents age 16-39 and all respondents age 40 and over, respectively. The proportion for the 16-39 year old stratum (w^1) was 53.0 percent while the proportion for the 40 and over stratum (w^2) was 47.0 percent.

As shown in Table 57, the disproportionate sampling increases the confidence interval by an average of 0.7 percent, compared to a simple random sample of the same size. This means the sample design slightly decreases the sampling precision for total population estimates, while increasing the precision of sampling estimates for the sub-sample aged 16-39 years old. Since the average difference in the confidence interval between the stratified disproportionate sample and a simple random sample is less than one percentage point, the sampling error table for a simple random sample will provide a reasonable approximation of the precision of sampling estimates in the survey.

TABLE 57
Design Effect On Confidence Intervals For Sample Estimates Between Disproportionate Sample Used In Occupant Protection Survey And A Proportionate Sample Of Same Size

VARIABLE (Version 1 only)	p=	------------------- CONFIDENCE INTERVALS -------------------- PERCENTAGE POINTS ± AT 95% CONFIDENCE LEVEL: HYPOTHETICAL PROPORTIONATE SAMPLING*	CURRENT DIS-PROPORTIONATE SAMPLING	DIFFERENCE IN CONFIDENCE INTERVALS ABOUT ESTIMATES
Driven in the past year	*89.2%*	*0.77*	*0.78*	*1.3%*
Drunk alcohol in past year	*63.4%*	*1.21*	*1.23*	*1.7%*
Always use safety belt (N=5502)	*85.1%*	*0.94*	*0.94*	----
Dislike safety belts (N=5505)	*33.1%*	*1.24*	*1.26*	*1.6%*
Always use passenger belt (N=5655)	*82.7%*	*0.98*	*0.98*	----
Favor (a lot) safety belt laws	*69.3%*	*1.15*	*1.16*	*.9%*
Should be primary enforcement	*63.9%*	*1.20*	*1.22*	*.9%*
Ever ticketed by police for seatbelt	*9.3%*	*0.73*	*0.72*	*-1.4%*
Ever injured in vehicle accident	*23.6%*	*1.06*	*1.08*	*1.9%*
Drives a car for work almost every day	*17.2%*	*0.94*	*0.96*	*2.1%*
Set a good example for others (N=5413) (reason for using safety belts)	*74.1%*	*1.17*	*1.19*	*1.7%*
Driver-side air bag in vehicle (N=5551)	*76.5%*	*1.12*	*1.14*	*1.8%*
Race: Black/African American	*8.6%*	*0.70*	*0.70*	----
Ethnicity: Hispanic	*13.2%*	*0.84*	*0.81*	*-3.6%*
Gender: Male	*48.0%*	*1.24*	*1.27*	*2.4%*
AVERAGE DIFFERENCE IN CONFIDENCE INTERVALS				*0.7%*

* Total sample proportions using SRS formula
Unless specified otherwise N=6180

Estimating Statistical Significance

The estimates of sampling precision presented in the previous section yield confidence bands around the sample estimates, within which the true population value should lie. This type of sampling estimate is appropriate when the goal of the research is to estimate a population distribution parameter. However, the purpose of some surveys is to provide a comparison of population parameters estimated from independent samples (e.g. annual tracking surveys) or between subsets of the same sample. In such instances, the question is not simply whether or not there is any difference in the sample statistics that estimate the population parameter, but rather is the difference between the sample estimates statistically significant (i.e., beyond the expected limits of sampling error for both sample estimates).

To test whether or not a difference between two sample proportions is statistically significant, a rather simple calculation can be made. The maximum expected sampling error (i.e., confidence interval in the previous formula) of the first sample is designated ***s1*** and the maximum expected sampling error of the second sample is ***s2***. The sampling error of the difference between these estimates is ***sd*** and is calculated as:

$$\mathrm{sd} = \sqrt{(s1^2 + s2^2)}$$

Any difference between observed proportions that exceeds ***sd*** is a statistically significant difference at the specified confidence interval. Note that this technique is mathematically equivalent to generating standardized tests of the difference between proportions.

An illustration of the pooled sampling error between sub-samples for various sizes is presented in Table 58. This table can be used to determine the size of the difference in proportions between drivers and non-drivers or other sub-samples that would be statistically significant.

TABLE 58. Pooled Sampling Error Expressed As Percentages For Given Sample Sizes (Assuming P=Q)

Sample Size																	
4000	14.1	10.0	7.1	5.9	5.1	4.7	4.3	4.0	3.8	3.6	3.5	3.0	2.7	2.5	2.4	2.3	2.2
3500	14.1	10.0	7.1	5.9	5.2	4.7	4.3	4.1	3.8	3.7	3.5	3.0	2.7	2.6	2.4	2.3	
3000	14.1	10.0	7.2	5.9	5.2	4.7	4.4	4.1	3.9	3.7	3.6	3.1	2,8	2.7	2.5		
2500	14.1	10.0	7.2	6.0	5.3	4.8	4.5	4.2	4.0	3.8	3.7	3.2	2.9	2.8			
2003	14.2	10.1	7.3	6.1	5.4	4.9	4.6	4.3	4.1	3.9	3.8	3.3	3.1				
1500	14.2	10.2	7.4	6.2	5.5	5.1	4.7	4.5	4.3	4.1	4.0	3.6					
1000	14.3	10.3	7.6	6.5	5.8	5.4	5.1	4.8	4.7	4.5	4.4						
900	14.4	10.4	7.7	6.5	5.9	5.5	5.2	4.9	4.8	4.6							
800	14.4	10.4	7.8	6.6	6.0	5.6	5.3	5.1	4.9								
700	14.5	10.5	7.9	6.8	6.1	5.7	5.5	5.2									
600	14.6	10.6	8.0	6.9	6.3	5.9	5.7										
500	14.7	10.8	8.2	7.2	6.6	6.2											
400	14.8	11.0	8.5	7.5	6.9												
300	15.1	11.4	9.0	8.0													
200	15.6	12.1	9.8														
100	17.1	13.9															
50	19.8																
	50	100	200	300	400	500	600	700	800	900	1000	1500	2003	2500	3000	3500	4000
								Sample Size									

2003 SURVEY RESULTS

APPENDIX B: STATE HIGHWAY SAFETY LAWS AT TIME OF SURVEY

TABLE 59
Key Provisions Of State Highway Safety Laws At Time Of Survey

STATE	Enforcement	FINE	Seating Positions Covered
ALABAMA	Primary	$25	Front
ALASKA	Secondary	$15	All
ARIZONA	Secondary	$10	All
ARKANSAS	Secondary	$25	Front
CALIFORNIA	Primary	$20	All
COLORADO	Secondary	$15	Front
CONNECTICUT	Primary	$37	Front
DELAWARE	Secondary	$20	Front
DIST. OF COLUMBIA	Primary	$50	All
FLORIDA	Secondary	$30	Front
GEORGIA	Primary	$15	Front
HAWAII	Primary	$45	Front
IDAHO	Secondary	$5	Front
ILLINOIS	Secondary	$25	Front
INDIANA	Primary	$25	Front
IOWA	Primary	$25	Front
KANSAS	Secondary	$10	Front
KENTUCKY	Secondary	$25	All
LOUISIANA	Primary	$25	Front
MAINE	Secondary	$25-$50	All
MARYLAND	Primary	$25	Front
MASSACHUSETTS	Secondary	$25	All
MICHIGAN	Primary	$25	Front
MINNESOTA	Secondary	$25	Front
MISSISSIPPI	Secondary	$25	Front
MISSOURI	Secondary	$10	Front
MONTANA	Secondary	$20	All
NEBRASKA	Secondary	$25	Front
NEVADA	Secondary	$25	All
NEW HAMPSHIRE	No law		
NEW JERSEY	Primary	$20	Front
NEW MEXICO	Primary	$25	All
NEW YORK	Primary	$50-$100	Front
NORTH CAROLINA	Primary	$25	Front
NORTH DAKOTA	Secondary	$20	Front
OHIO	Secondary	$25	Front
OKLAHOMA	Primary	$20	Front
OREGON	Primary	$75	All
PENNSYLVANIA	Secondary	$10	Front
RHODE ISLAND	Secondary	$50	All
SOUTH CAROLINA	Secondary	$10	All
SOUTH DAKOTA	Secondary	$20	Front
TENNESSEE	Secondary	$10	Front
TEXAS	Primary	$25-$50	Front
UTAH	Secondary	$45	All
VERMONT	Secondary	$10	All
VIRGINIA	Secondary	$25	Front
WASHINGTON	Primary	$86	All
WEST VIRGINIA	Secondary	$25	Front
WISCONSIN	Secondary	$10	All
WYOMING	Secondary	$25	All

DOT HS 809 789
September 2004

U.S. Department
of Transportation
National Highway
Traffic Safety
Administration

NHTSA
People Saving People
www.nhtsa.dot.gov