A SAMPLE SURVEY DESIGN TO ESTIMATE THE MEANS OF TWO OVERLAPPING SUBPOPULATIONS

Timothy C. Coburn and William D. Warde, Oklahoma State University

1. Introduction

A common situation in survey sampling is for several organizations to collect information on a regular basis from the same segment of a population. In addition, some of the same units may actually be selected for use in two or more surveys, and the information to be obtained from them may be almost identical. When much of the required data for all the surveys could be collected simultaneously from the same set of sampling units, or from subunits of those units, this practice is statistically and cost inefficient.

The procedure of simultaneously collecting data on the same unit for several surveys is sometimes referred to as integrated survey sampling. When a single population is composed of two or more overlapping subpopulations, independent surveys of the subpopulations can be integrated and the sample designs can be modified to either (1) reduce or eliminate multiple coverage in the overlap domain, or (2) to improve the estimates of the parameters of the individual subpopulations by advantageously combining information available from all the surveys. These features resemble some of the features of the multiple frames techniques. However, while the multiple frames approach addresses the question of how to best sample a single population covered by two or more distinct, but overlapping, lists of units (frames), the integrated survey approach deals with the problems of sampling two or more distinct, but overlapping, populations or subpopulations, each covered by a single frame or list of units.

This investigation addresses the general problems of integrated survey sampling and estimation in the overlapping subpopulations context. For the special case of a single population covered by exactly two overlapping subpopulations, a basic integrated sample survey design is proposed which yields a more precise estimator of the means of the two individual subpopulations (combining all the available survey data) than estimators obtained with conventional survey designs. Alternatively, for fixed levels of precision, procedures are described for reducing the sample sizes of the two surveys in order to minimize coverage of the overlap domain, resulting in a more cost efficient system of surveys.

2. An Example of Overlapping Subpopulations

An example of overlapping subpopulations is found in the federal welfare system. The federal welfare system is most generally defined as the collection of federally funded family nutritional/income support programs, the agencies and their staffs which administer and govern those programs, and the population of all family units receiving at least one program benefit. Although there are a number of these programs in existence serving specific interests throughout the country, there are three major programs which collectively support the majority of "welfare recipients" nationwide. It is these three programs which are of specific interest: Aid to Families with Dependent Children (AFDC), the basic welfare grant program administered by the Health Care Financing Administration (HCFA) of the U.S. Department of Health, Education, and Welfare (DHHS); the Food Stamp Program (FSP), administered by the Food and Nutrition Service (FNS) of the U.S. Department of Agriculture (USDA); and the Medicaid program (Med), administered by the Social Security Administration (SSA) of DHHS.

Every state directs these three programs within its geographic boundaries under federal statutes and with the support of public funds. In most cases the programs are managed independently of one another, either by different state agencies or by different staffs within the same agency, and each program supports its own participating constituency. In this regard the population of welfare recipients in each state actually consists of three uniquely defined and uniquely governed subpopulations. Under the collective regulations currently in force, it is possible, and highly likely, that a family unit eligible to participate in one program may also be eligible to participate in one or both of the other programs. Consequently, the subpopulations, though uniquely defined, inherently overlap. A schematic of this general overlapping subpopulations situation is provided in the following figure.

Figure 1. Overlapping Subpopulations in the Federal Welfare System

Every six months each state is required to conduct a quality control survey of a sample of family units residing in the state and participating in each of the three major welfare programs. There are three independent samples of fixed size from which to collect and analyze data (in most cases this requires the expertise of three separate, trained staffs). The purpose of these surveys is to vali-
date the management practices of the state agencies
directing the programs by determining the number of
participating family units obtaining benefits in
error. Of particular interest is the proportion of
family units certified to participate in a given
program, but which, because of oversight or fraud
by the caseworker or recipient, are totally ineli-
gible. States are subject to fiscal sanctions if
their "ineligibility rates" in any of the three
programs exceed established tolerances. Conse-
sequently there is considerable interest in integ-
rate these sample surveys to reduce overall costs
and/or to improve the estimates of the individual
subpopulation characteristics of interest.

3. A Basic Overlapping Sample Surveys Design
Suppose a population of size N is composed of
two overlapping subpopulations of sizes N and M,
respectively; and suppose two independent sample
surveys are conducted over the population, with
each survey aimed at a particular subpopulation.
The staffs of two distinct agencies or organiza-
tions conduct the surveys. Let the survey of sub-
population 1 (having size N) be designated as the
primary survey (primary subpopulation, primary sam-
ping, etc.). Let the overlap domain be of size N2
(= M2), so that N = N1 + N2 and M = M1 + M2.

Although the subpopulations are known to ovel-
lap, it cannot be known prior to sampling which
population elements fall in the overlap domain.
Assume the sampling units for both surveys are of
the same size (i.e., same definition of sampling unit),
and that both surveys obtain identical measurements
on the units for the characteristics of interest.

Simple random samples of fixed size n and m,
respectively, are selected for the two surveys. The
units selected in each sample fall into two catego-
ries, or strata—those which belong to the overlap
domain and those which do not (called "mixed" and
"non-mixed" units, respectively. Assume no du-
plicate units are selected in the two samples owing
to sampling the overlap domain twice.

The surveys are conducted by the two organiza-
tions. Within the context of each survey it is
first determined for each sampled unit whether or
not it falls in the overlap domain. If a sample
unit is determined to be a "mixed" unit, it is
first surveyed with respect to membership in the
subpopulation from which it was selected. Within
the scope of this investigation, or subsequent to
it, additional information is obtained on the char-
acteristics of interest to the other survey. After
the two surveys are completed, each of the two ori-
ginal samples is post-stratified into the two catego-
ries—"mixed" and "non-mixed" units. In this
manner four subsamples of random size are formed:
N1 "non-mixed" units (with respect to the primary
subpopulation) and N2 "mixed" units in the primary
sample; and M1 "non-mixed" units (with respect to
the second subpopulation) and M2 "mixed" units in
the second sample. Note that n = n1 + n2 and
m = m1 + m2 (n1, m1, i = 1,2 all non-zero).

The two subsamples of "mixed" units are two independent
samples from the overlap domain. Improved es-
timates for the characteristics of interest in each
subpopulation can be computed by advantageously
combining the information available in the four
subsamples.

4. Estimating the Mean of Either Subpopulation
Assume the total size of the overlap domain of
the two subpopulations is known. Let N = N1/N
and M = M1/M, N + M = N, be the usual weights
appropriate to stratified sampling. Let y be a
characteristic of interest in the primary survey,
and let y1hi be the value of y on the ith unit in
stratum h of the primary sample (h = 1,2). Then
an unbiased estimate of y1, the true mean of the
primary subpopulation, obtained via proportional
stratification of a single sample of size n, is given by

\[ \bar{y}_{11}^* = \frac{1}{n} \sum_{i=1}^{n} y_{1hi} \]  

Var(\bar{y}_{11}^*) = \frac{\sum_{h=1}^{2} \text{Var}(y_{1hi})}{n} = \sum_{h=1}^{2} \frac{1}{n} \text{Var}(y_{1hi}) 

where \( \text{Var}(y_{1hi}) \) is the variance of the characteristic
y in stratum h of the primary sample. Let \( \bar{y}_{11} \) be the
mean of the sample of size n from the primary
subpopulation, and let \( \bar{y}_{21} \) be the mean of the
sample of size m from the second subsample.

The average value of \( \text{Var}(\bar{y}_{11}^*) \) over all possible
non-zero n must be obtained. Ignoring the case
n = 0, the expression can be given for \( \text{Var}(\bar{y}_{11}^*) \)

Var(\bar{y}_{11}^*) = \frac{n}{n-1} \text{Var}(\bar{y}_{11}) + \frac{1}{n} \text{Var}(\bar{y}_{21}) 

where \( \text{Var}(\bar{y}_{11}) \) and \( \text{Var}(\bar{y}_{21}) \) are
the variances of the primary and secondary es-
timates, respectively. A two-sample estimate of the
mean, \( \bar{y}_{11}^* \), of the primary subpopulation for the char-
acteristic y is given by

\[ \bar{y}_{11}^* = \frac{1}{n} \sum_{i=1}^{n} y_{1hi} + \frac{1}{m} \sum_{j=1}^{m} y_{2ji} \]

Now suppose additional information via a sam-
ple from the second subpopulation is available on
stratum 2 (overlap domain). Let y1hi and y2hi be
values of the characteristic y obtained on the
ith units in strata h from subpopulations 1 and 2,
respectively. A two-sample estimate of the mean,
\( \bar{y}_{11}^* \), of the primary subpopulation for the char-
acteristic y is given by

\[ \bar{y}_{11}^* = \frac{1}{n} \sum_{i=1}^{n} y_{1hi} + \frac{1}{m} \sum_{j=1}^{m} y_{2ji} \]

where \( \bar{y}_{11} \) and \( \bar{y}_{21} \) are unbiased for \( \bar{y}_{11} \) and \( \bar{y}_{21} \),
respectively. The unbiased estimate of the mean of the
hth strata in subpopulations 1 and 2, respectively
(same expression can be given for y2h), is found
by

\[ \bar{y}_{1h} = \frac{1}{n} \sum_{i=1}^{n} y_{1hi} \]

\[ \bar{y}_{2h} = \frac{1}{m} \sum_{j=1}^{m} y_{2hi} \]
\[ \text{Var}(\bar{y}^*_2|\beta) = \frac{n_1^2}{n_1} + \frac{n_2^2}{n_2} + \frac{V_2^2(1-\beta)^2}{V_2^2}. \]

Now \( \text{Var}(\bar{y}^*_2|\beta) \) is minimum when \( \beta = \frac{n_2}{n_2 + m_2} \). Making this substitution in the above equation,

\[ \text{Var}(\bar{y}^*_2|\beta) = \frac{n_2}{n_2 + m_2} + \frac{V_2^2(1-\beta)^2}{V_2^2}. \]

The average value of \( \text{Var}(\bar{y}^*_2|\beta) \) over all possible non-zero values of \( n_1, n_2, \) and \( m_2 \) must now be obtained.

\[ \text{Var}(\bar{y}^*_2) = \mathbb{E}_{n_2,m_2}[\text{Var}(\bar{y}^*_2|\beta)] = \frac{V_1^2 V_2^2}{n_2} + \frac{V_1^2 V_2^2}{V_1^2} + \frac{V_1^2 V_2^2}{(1+\Delta)^2}, \]

where \( \Delta = \frac{V_1^2 V_2^2}{n_2}, \mathbb{E}[\frac{1}{n_2}] = \frac{1}{V_1^2 + V_2^2} + \frac{n_{11} V_1^2 + V_2^2}{(V_1^2 + V_2^2)^2}, \)

and \( V_1 = n_1/\pi, V_2 = n_2/\pi \). Analogous algebraic expressions may be obtained for the variances of multi-sample estimators of other subpopulation parameters.

5. Comparison with Conventional Estimators

To the order of approximation used in computing the variance, it can be shown that \( \bar{y}^*_2 \) is uniformly more precise than \( \bar{y}^*_1 \); the estimate of the mean of the primary subpopulation obtained with conventional single sample post-stratification. In addition, if

\[ \frac{S_1^2}{S_2^2} < \frac{V_1}{V_2}, \]

then \( \bar{y}^*_2 \) is more precise than \( \bar{y}^*_1 \) obtained with proportional stratification. An approximate condition is given by

\[ S_1^2/S_2^2 < n. \]

The relative precision of \( \bar{y}^*_1 \) to \( \bar{y}^*_2 \) was computed for each of a number of combinations of values of the parameters \( n_1, m_1, V_1^2, \) and \( V_2^2 \) in the expression for the variance of \( \bar{y}^*_1 \), and for three different combinations of the within-stratum variances, \( S_1^2 \) and \( S_2^2 \). Some of the results are shown in Tables I-III.

6. Two Sample Size Reduction Schemes Based on the Two-Sample Estimator

Given two overlapping subpopulations, the size of the sample for the survey of the primary subpopulation may be reduced without altering the sample size for the second survey by solving for \( n' \) in the equation

\[ \text{Var}(\bar{y}^*_1|n',m,W_2,V_2) = C. \]

In this equation \( C \) is the desired precision of the estimate \( \bar{y}^*_1 \), \( n' \) is the reduced primary sample size, and \( m \) is the size of the sample for the second survey. Table IV displays the sizes of \( n' \) to which \( n \) may be reduced for several combinations of values of the parameters \( n, m, W_2, \) and \( V_2 \), when \( C \) is taken to be the precision associated with conventional single-sample post-stratification (of a primary sample size \( n \)).

It is also possible to reduce both original sample sizes by simultaneously solving the two equations

\[ \text{Var}(\bar{y}^*_1|n',m,W_2,V_2) = C \]

and

\[ \text{Var}(\bar{y}^*_2|n',m,W_2,V_2) = D. \]

In these equations, \( \bar{y}^*_2 \) is the two-sample estimate of the mean of the first subpopulation, \( \bar{y}^*_2 \) is the two-sample estimate of the mean of the second subpopulation, \( n' \) and \( m' \) are the sizes to which \( n \) and \( m \) may be reduced, respectively, \( C \) is the desired precision for \( \bar{y}^*_1 \), and \( D \) is the desired precision for \( \bar{y}^*_2 \). Table V shows some pairs of reduced sample sizes obtained when \( C \) and \( D \) are taken to be the precision associated with estimating the means of the two subpopulations using conventional single-sample post-stratification (of samples of sizes \( n \) and \( m \), respectively).

The second procedure described above yields a uniformly smaller combined total sample size for the surveys of two overlapping subpopulations than does the first procedure. Both procedures allow fewer sample units to be physically surveyed than would ordinarily be required, using conventional single-sample survey designs, to maintain a desired level of precision for estimates of the subpopulation parameters of interest. Accomplishment of this goal depends on the sharing of information among survey organizations as specified in the survey design previously described. Specific sample survey situations will dictate a choice between the two procedures.

7. Summary of Findings

In the context of two overlapping surveys, an estimate of the mean of either subpopulation, \( \bar{y}^*_2 \), can be obtained by combining information in samples of size \( n \) and \( m \), respectively, selected from the two subpopulations. \( \bar{y}^*_2 \), the estimate of the mean of the subpopulation of primary interest, is always more precise than \( \bar{y}^*_1 \), the estimate of the mean obtained by post-stratifying the single sample of size \( n \). The difference in the precision of \( \bar{y}^*_2 \) and the precision of \( \bar{y}^*_1 \) is greatest for any combination of stratum variances when \( m \) is large relative to the size of \( n \) and \( W_2 \) and \( V_2 \) are both large and about the same size. Even larger gains in precision are to be obtained using \( \bar{y}^*_2 \) if \( S_2^2 > S_1^2 \).

Therefore, if (1) the two subpopulations are about the same size and are substantially overlapped, (2) the size of the sample from the second subpopulation exceeds the size of the primary sample, and (3) the overlap domain is the most variable stratum, then \( \bar{y}^*_2 \) should be used to estimate the primary subpopulation mean. Otherwise \( \bar{y}^*_1 \) is about as precise as \( \bar{y}^*_2 \), and the conventional single-sample estimator is recommended if the additional administrative costs of operating in the overlapping surveys mode are substantial (analogous re-
results can be obtained regardless which subpopulation is chosen as the primary subpopulation. It can be demonstrated that these recommendations apply for small values of \( n \) and \( m \) as well as for large values, though the results noted above are likely to be more pronounced when \( n \) and \( m \) are both relatively large (for example, 100 or larger). In addition, for almost all choices of \( n, m, \frac{1}{\sqrt{n}}, \frac{1}{\sqrt{m}}, S_1^2, \) and \( S_2^2, \frac{Y^{**}_1}{Y^{**}_1} \) is also more precise than \( \frac{Y_1}{Y_1} \), the estimate of the mean obtained by proportional allocation of the single sample of size \( n \) among the strata of the primary subpopulation.

If the precision of the two-sample estimator, \( \frac{Y^{**}_1}{Y^{**}_1} \), is pre-specified, then the combined total sample size for the surveys of the two overlapping subpopulations can be reduced using either of two different procedures. The survey situation may allow only one of the sample sizes to be reduced, while the other remains fixed; or it may allow both sample sizes to be reduced simultaneously. The percent reduction in combined total sample size is greatest when both sample sizes are reduced simultaneously.

*This research was supported in part by the U.S. Department of Agriculture, Food and Nutrition Service, Grant No. 40-3198-9-888.

### TABLE I

**Relative Precision of \( \frac{Y^{**}_1}{Y^{**}_1} \) to \( \frac{Y_1}{Y_1} \), Large Sample Sizes, \( S_2 = S_1 \)**

<table>
<thead>
<tr>
<th>( n = 3m )</th>
<th>( n = 2m )</th>
<th>( n = m )</th>
<th>( n = \frac{3m}{2} )</th>
<th>( n = \frac{2m}{3} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{Y^{**}_1}{Y_1} )</td>
<td>( \frac{Y^{**}_1}{Y_1} )</td>
<td>( \frac{Y^{**}_1}{Y_1} )</td>
<td>( \frac{Y^{**}_1}{Y_1} )</td>
<td>( \frac{Y^{**}_1}{Y_1} )</td>
</tr>
<tr>
<td>( .2 )</td>
<td>( .4 )</td>
<td>( .6 )</td>
<td>( .8 )</td>
<td>( .2 )</td>
</tr>
<tr>
<td>1.11</td>
<td>1.21</td>
<td>1.22</td>
<td>1.12</td>
<td>1.15</td>
</tr>
<tr>
<td>1.31</td>
<td>1.43</td>
<td>1.46</td>
<td>1.52</td>
<td>1.25</td>
</tr>
<tr>
<td>1.42</td>
<td>1.66</td>
<td>1.81</td>
<td>1.92</td>
<td>1.31</td>
</tr>
<tr>
<td>1.51</td>
<td>1.91</td>
<td>2.23</td>
<td>2.49</td>
<td>1.56</td>
</tr>
</tbody>
</table>

### TABLE II

**Relative Precision of \( \frac{Y^{**}_1}{Y_1} \) to \( \frac{Y^{**}_1}{Y^{**}_1} \), Large Sample Sizes, \( S_2 = 25_1 \)**

<table>
<thead>
<tr>
<th>( n = 3m )</th>
<th>( n = 2m )</th>
<th>( n = m )</th>
<th>( n = \frac{3m}{2} )</th>
<th>( n = \frac{2m}{3} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{Y^{**}_1}{Y_1} )</td>
<td>( \frac{Y^{**}_1}{Y_1} )</td>
<td>( \frac{Y^{**}_1}{Y_1} )</td>
<td>( \frac{Y^{**}_1}{Y_1} )</td>
<td>( \frac{Y^{**}_1}{Y_1} )</td>
</tr>
<tr>
<td>( .2 )</td>
<td>( .4 )</td>
<td>( .6 )</td>
<td>( .8 )</td>
<td>( .2 )</td>
</tr>
<tr>
<td>1.04</td>
<td>1.14</td>
<td>1.21</td>
<td>1.34</td>
<td>1.20</td>
</tr>
<tr>
<td>1.52</td>
<td>1.75</td>
<td>1.97</td>
<td>2.16</td>
<td>1.40</td>
</tr>
<tr>
<td>1.60</td>
<td>1.99</td>
<td>2.38</td>
<td>2.60</td>
<td>1.42</td>
</tr>
<tr>
<td>1.61</td>
<td>2.13</td>
<td>2.59</td>
<td>2.99</td>
<td>1.42</td>
</tr>
</tbody>
</table>

### TABLE III

**Relative Precision of \( \frac{Y^{**}_1}{Y^{**}_1} \) to \( \frac{Y_1}{Y_1} \), Large Sample Sizes, \( S_2 = \frac{1}{2}S_1 \)**

<table>
<thead>
<tr>
<th>( n = 3m )</th>
<th>( n = 2m )</th>
<th>( n = m )</th>
<th>( n = \frac{3m}{2} )</th>
<th>( n = \frac{2m}{3} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{Y^{**}_1}{Y_1} )</td>
<td>( \frac{Y^{**}_1}{Y_1} )</td>
<td>( \frac{Y^{**}_1}{Y_1} )</td>
<td>( \frac{Y^{**}_1}{Y_1} )</td>
<td>( \frac{Y^{**}_1}{Y_1} )</td>
</tr>
<tr>
<td>( .2 )</td>
<td>( .4 )</td>
<td>( .6 )</td>
<td>( .8 )</td>
<td>( .2 )</td>
</tr>
<tr>
<td>1.09</td>
<td>1.10</td>
<td>1.11</td>
<td>1.14</td>
<td>1.06</td>
</tr>
<tr>
<td>1.17</td>
<td>1.22</td>
<td>1.25</td>
<td>1.28</td>
<td>1.14</td>
</tr>
<tr>
<td>1.26</td>
<td>1.38</td>
<td>1.50</td>
<td>1.54</td>
<td>1.20</td>
</tr>
<tr>
<td>1.37</td>
<td>1.63</td>
<td>1.81</td>
<td>1.99</td>
<td>1.28</td>
</tr>
</tbody>
</table>
### TABLE IV

Reduced Primary Sample Sizes for the Surveys of Two Overlapping Subpopulations

<table>
<thead>
<tr>
<th>n</th>
<th>n</th>
<th>$S_2 = \frac{s_2^2}{s_1^2}$</th>
<th>$S_2 = \frac{z_2^2}{z_1^2}$</th>
<th>$S_2 = \frac{1}{z_1^2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>200</td>
<td>178 171 168 166 162 158 154</td>
<td>162 150 145 142 138 134 130</td>
<td>188 184 180 176 172 168 164</td>
</tr>
<tr>
<td>400</td>
<td>200</td>
<td>170 154 146 140 155 139 133</td>
<td>159 137 125 119 113 107 100</td>
<td>182 178 174 170 166 162 158</td>
</tr>
<tr>
<td>1000</td>
<td>200</td>
<td>164 148 136 130 140 125 119</td>
<td>158 136 121 113 107 100 94</td>
<td>167 157 147 137 128 120 111</td>
</tr>
<tr>
<td>1400</td>
<td>200</td>
<td>160 144 132 126 136 122 116</td>
<td>154 132 117 110 104 98 91</td>
<td>154 144 134 124 115 106 97</td>
</tr>
<tr>
<td>1800</td>
<td>200</td>
<td>156 140 128 122 132 118 112</td>
<td>148 126 111 104 98 91 84</td>
<td>144 134 124 114 105 96 87</td>
</tr>
<tr>
<td>2200</td>
<td>200</td>
<td>152 136 122 116 126 112 106</td>
<td>146 124 109 102 95 88 80</td>
<td>134 124 114 104 95 86 77</td>
</tr>
<tr>
<td>2600</td>
<td>200</td>
<td>148 132 118 112 122 108 102</td>
<td>136 114 100 93 86 79 70</td>
<td>124 114 104 94 85 76 67</td>
</tr>
<tr>
<td>3000</td>
<td>200</td>
<td>144 128 114 108 118 104 98</td>
<td>132 110 96 89 82 73 64</td>
<td>114 104 94 84 75 66 57</td>
</tr>
<tr>
<td>3400</td>
<td>200</td>
<td>140 124 108 102 112 98 90</td>
<td>124 102 88 81 72 63 64</td>
<td>104 94 84 74 65 56 47</td>
</tr>
<tr>
<td>3800</td>
<td>200</td>
<td>136 120 104 98 108 94 86</td>
<td>116 94 80 72 63 54 45</td>
<td>94 84 74 64 55 46 36</td>
</tr>
<tr>
<td>4200</td>
<td>200</td>
<td>132 116 100 92 100 86 78</td>
<td>106 84 68 60 52 42 33</td>
<td>84 74 64 54 45 36 26</td>
</tr>
<tr>
<td>4600</td>
<td>200</td>
<td>128 112 92 88 92 78 70</td>
<td>96 72 56 48 40 30 20</td>
<td>74 64 54 44 34 24 14</td>
</tr>
<tr>
<td>5000</td>
<td>200</td>
<td>124 108 96 92 96 80 72</td>
<td>86 62 44 36 28 18 8</td>
<td>64 54 44 34 24 14 8</td>
</tr>
<tr>
<td>5400</td>
<td>200</td>
<td>120 102 90 90 94 80 72</td>
<td>76 52 34 26 18 10 0</td>
<td>54 44 34 24 14 8 0</td>
</tr>
<tr>
<td>5800</td>
<td>200</td>
<td>116 100 84 84 88 76 68</td>
<td>66 40 22 14 8 0 0</td>
<td>44 34 24 14 8 0 0</td>
</tr>
<tr>
<td>6200</td>
<td>200</td>
<td>112 94 80 80 84 70 62</td>
<td>56 30 14 4 0 0 0</td>
<td>34 24 14 8 0 0 0</td>
</tr>
<tr>
<td>6600</td>
<td>200</td>
<td>108 92 76 76 78 68 60</td>
<td>46 22 0 0 0 0 0</td>
<td>24 14 8 0 0 0 0</td>
</tr>
<tr>
<td>7000</td>
<td>200</td>
<td>104 88 74 74 76 62 58</td>
<td>36 14 0 0 0 0 0</td>
<td>14 8 0 0 0 0 0</td>
</tr>
<tr>
<td>7400</td>
<td>200</td>
<td>100 84 70 70 70 60 56</td>
<td>26 0 0 0 0 0 0</td>
<td>8 0 0 0 0 0 0</td>
</tr>
</tbody>
</table>

192
### TABLE V

Reduced Sample Sizes for the Surveys of Two Overlapping Subpopulations

<table>
<thead>
<tr>
<th>(a)</th>
<th>(m)</th>
<th>(n)</th>
<th>(n_{2})</th>
<th>(n_{3})</th>
<th>(n_{4})</th>
<th>(n_{5})</th>
<th>(n_{6})</th>
<th>(n_{7})</th>
<th>(n_{8})</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>200</td>
<td>180</td>
<td>180</td>
<td>173</td>
<td>173</td>
<td>167</td>
<td>170</td>
<td>170</td>
<td>167</td>
</tr>
<tr>
<td>4</td>
<td>173</td>
<td>173</td>
<td>160</td>
<td>160</td>
<td>152</td>
<td>152</td>
<td>146</td>
<td>147</td>
<td>147</td>
</tr>
<tr>
<td>6</td>
<td>169</td>
<td>169</td>
<td>152</td>
<td>152</td>
<td>140</td>
<td>140</td>
<td>132</td>
<td>132</td>
<td>132</td>
</tr>
<tr>
<td>8</td>
<td>169</td>
<td>169</td>
<td>147</td>
<td>147</td>
<td>132</td>
<td>132</td>
<td>121</td>
<td>121</td>
<td>121</td>
</tr>
<tr>
<td>400</td>
<td>400</td>
<td>172</td>
<td>172</td>
<td>167</td>
<td>167</td>
<td>160</td>
<td>160</td>
<td>152</td>
<td>152</td>
</tr>
<tr>
<td>4</td>
<td>177</td>
<td>177</td>
<td>160</td>
<td>160</td>
<td>152</td>
<td>152</td>
<td>146</td>
<td>146</td>
<td>146</td>
</tr>
<tr>
<td>6</td>
<td>169</td>
<td>169</td>
<td>152</td>
<td>152</td>
<td>140</td>
<td>140</td>
<td>132</td>
<td>132</td>
<td>132</td>
</tr>
<tr>
<td>8</td>
<td>169</td>
<td>169</td>
<td>147</td>
<td>147</td>
<td>132</td>
<td>132</td>
<td>121</td>
<td>121</td>
<td>121</td>
</tr>
<tr>
<td>600</td>
<td>600</td>
<td>172</td>
<td>172</td>
<td>167</td>
<td>167</td>
<td>160</td>
<td>160</td>
<td>152</td>
<td>152</td>
</tr>
<tr>
<td>4</td>
<td>177</td>
<td>177</td>
<td>160</td>
<td>160</td>
<td>152</td>
<td>152</td>
<td>146</td>
<td>146</td>
<td>146</td>
</tr>
<tr>
<td>6</td>
<td>169</td>
<td>169</td>
<td>152</td>
<td>152</td>
<td>140</td>
<td>140</td>
<td>132</td>
<td>132</td>
<td>132</td>
</tr>
<tr>
<td>8</td>
<td>169</td>
<td>169</td>
<td>147</td>
<td>147</td>
<td>132</td>
<td>132</td>
<td>121</td>
<td>121</td>
<td>121</td>
</tr>
</tbody>
</table>