TELEPHONE SAMPLE DESIGNS FOR THE BLACK HOUSEHOLD POPULATION

Kathryn M. Inglis, Robert M. Groves, Steven G. Heeringa, The University of Michigan

1. Introduction

Sample surveys of rare populations,where no frame for the rare population exists, often have large per unit costs relative to similar designs for the full population. This is especially true in household surveys of demographic subgroups for which screening is used to locate eligible sample units,such as for the Black population. In recent years in the United States, telephone sampling methods have been proposed as cost efficient tools for sampling and interviewing rare populations. In this paper we examine a variety of telephone sample designs for the U.S. Black household population, which forms approximately ten percent of the full population.

The use of any survey design must be evaluated based on its costs and errors relative to feasible alternatives. The use of telephone sampling and interviewing implies that Blacks living in households without telephones are not covered by the survey procedures. Such persons tend to be poorer, and younger than those with telephones (Thornberry and Massey, 1983). The survey experiments described in this paper were part of a pre/post study of Black political attitude and electoral behavior in the 1984 presidential election. To the extent that Blacks without telephones have attitudes and voting behaviors that are different from those with telephones the survey estimates would differ from Black household population parameters. While not ignoring the noncoverage error associated with telephone survey designs for the Black population, this paper focuses on differential cost efficiencies and sampling error that might result from several telephone sample designs.

The survey designer has three principal tools to reduce the costs of screening:

a) stratification of units by proportion Black, and disproportionate allocation to high density Black strata;
b) use of measures of size for the Black population for the first stage units in a multistage design; and
c) increase in the number of selected elements per first stage unit .

Stratification of the telephone population by race attempts to isolate areas with high proportions of telephone subscribers who are Black. Large sampling fractions are then applied to those strata, relative to those with lower proportions Black. Thus, the total number of households that have to be contacted in order to obtain one interview with an eligible Black is smaller than that obtained with an epsem sample of the household population. Consequently,the screening costs for locating a sample of Black households is reduced. In telephone samples the basic geographical unit for stratification is the wire center or telephone exchange, to which one or more prefixes may be assigned. Unfortunately, no counts of the subscriber population by racial characteristics are available in general on these sampling units. Thus, proxy indicators of high density Black areas must be used. The experiments described in this paper examined the value of such proxy indicators.

Manipulation of cluster definition and PPS selection procedures, were explored in the context of a design proposed by Waksberg (1978), whereby a two stage cluster sample is chosen using rejection rules at the first and second stage. This design selects clusters of 100 consecutive telephone numbers with probabilities proportional to the number of working household numbers in the cluster. Second stage sampling of household numbers is performed with conditional probabilities of selection inversely proportional to the number of working household numbers in the cluster. Thereby, the design yields an epsem sample of household numbers, but clusters them so that the proportion of household numbers in the entire set is higher than that obtained by direct element sampling methods.

Blair and Czaja (1982) present a novel alteration of the Waksberg-Mitofsky technique, using rejection rules for subpopulations at both stages. For the Black population this method includes at the first stage only clusters whose primary number was assigned to a Black household and then samples Black household numbers within those clusters. This results in a sample of clusters chosen with probabilities proportional to the number of Black households in each cluster. Blair and Czaja found that the percentage of Black households among household numbers chosen increased from 9 percent for the first stage to 25 percent for the second stage numbers. Given the compensating probabilities of selection in the two stages, this design greatly reduces the level of screening required to obtain any given sample size of Black households. A similar alteration of measures of size was employed in the experiments described in this paper.

In the Blair and Czaja design some of the sample clusters had too few Black household numbers to yield the number of elements per cluster required (10 in their case) for an epsem design of Black households. In addition, relatively large screening costs are incurred at the first stage of selection for this design; over 44 numbers must be dialed to locate one Black household at the primary stage of sampling. The classical solution to these two problems is the selection of a smaller number of larger sized clusters (with larger numbers of sample elements per clusters).

The analyses reported here examined the use of cluster sizes of 100, 200, and 400 consecutive numbers each. The extension of the cluster definition beyond 100 consecutive numbers was suggested by observations about the assignment of numbers within prefixes. The following appears to be the most common pattern: 1) almost all household numbers within a prefix serve units located within the geographical boundaries of the exchange, 2) there is little geographical clustering of assignments within wire centers (i.e., neighbors do not tend to have consecutive telephone numbers, nor need they have numbers in the same prefix), 3) there is more diversity in the percentage of household numbers among 1000 series than among 100 series within the same 1000 bank of numbers. These impressions are the result of several years of telephone sampling at the Survey Research Center. Observations 1)−3) suggest that the expansion of the cluster definition from 100 consecutive numbers to a larger number might permit the use of larger numbers of secondary selections with little reduction in the proportion of those numbers which were Black households.

2. The Study Design

Several design alternatives were examined during a pilot phase of the survey which were not used in the production phase.

2.1 The Pilot Study

In preparation for the production study, a two wave pilot study was undertaken. For the first wave 1400 RDD primary numbers were systematically sampled from the 34,389 six-digit area code/central office code combinations listed on the current A.T. and T. tape. Prior to selection, three strata were defined: 1) exchanges corresponding to the central cities of large Standard Metropolitan Statistical Areas (e.g., Chicago city, for the Chicago SMSA), 2) other exchanges in selected southern states (Virginia, North Carolina, South Carolina, Florida, Georgia, Alabama, Mississippi, Louisiana), and 3) the balance of exchanges in the coterminous United States. An equal probability sample of primary numbers was selected from each of these three strata. For the second wave an additional 500 primary numbers were selected from the first and second strata only. A disproportionate allocation was used.

For the purposes of this paper, results from the two pilot study waves were pooled and weighted appropriately to correct for unequal selection probabilities.

Primary numbers in each of the Pilot study samples were dialed and screened for their Black household status. If the sampled number reached a Black household, it simultaneously identified three different sample clusters. As shown in Table 1, each number can be viewed as belonging to a single 100 series, a single 200 series, and a single 400 series. For example, the number 313-764-4424 is a member of the 4400-4499 hundred series, the 4400-4599 two hundred series, the 4400-4799 four hundred series. The pilot study sampled secondary numbers from each of these three sets of clusters. The second stage cluster sizes of Black households were set at 3, 6, and 9 respectively for the three sets of clusters excluding the primary number. In both the primary and secondary stages of selection, if the race of the household was not known then it was assumed to be a nonBlack household.

2.2 The Production Study

The production study used the same stratification , and implemented sampling fractions in the ratio 3:2:1 over the three strata, with 11,123 primary numbers selected overall. Clusters of 200 consecutive numbers were used with an expected cluster size of 5.5 (not counting the primary number). Primary and secondary stage rejection rules were identical to those used for the pilot study.

3. Results of the Pilot Study

Table 1 presents the disposition of the secondary numbers by cluster type and stratum. Of most interest is the proportion of secondary numbers assigned to Black households for the different cluster definitions. For the 100 series in the full population, .134 of all numbers are Black household numbers. This implies that .223 of the households sampled were Black, compared to the .25 Black households found by Blair and Czaja. For the 200 series clusters .124 of all numbers are Black household numbers. For the 400 series, .115 of all numbers are assigned to Blacks. All of these numbers are within sampling error of each other (the standard error of each estimate is at least .02). That is, there is no large decrease in the proportion eligible in moving from 100 to 400 consecutive numbers. These rates imply that while 100 series clusters on the average can support 13 or 14 sample Black households, the 400 series might support on the average cluster sizes of 46 sample Black households. The larger number of sample Black households permit the researcher to reduce radically sampling costs and screening costs.

Table 1 also permits comparison of the proportion of eligible secondary numbers for the three different strata used in the pilot study. For all the cluster definitions the same result applies —the large SMSA telephone exchanges offer close to a doubling of the eligibility rate versus the overall population (.21 versus .12 or .13). The medium density stratum, consisting of nonSMSA exchanges in selected Southern states, has eligibility rates below that of the nation as a whole (between .08 and .10). The low density stratum, the rest of the country, has similar rates (between .7 and .085). Since the high density stratum covers about thirty-six percent of the Black telephone household population, this stratification is an effective tool for cost reduction in screening.

4. Results of the Production Survey

It was judged most prudent to use the 200 series cluster definition for the final design. Table 2 presents the results from both the primary and secondary number screening. About 13 percent of all secondary numbers were Black households (the standard error about this is .6 percent). This compares to the 12 percent in the pilot study. The comparison of the results for the primary stage of selection with those of the secondary stage illustrate the large gains possible by using PPS two stage sampling for Black households. In the epsem selection of primary numbers only 2 percent of the numbers generated are Black household numbers; in the secondary selections 13 percent are. The relative gains through PPS selection are most dramatic in the low density Black stratum. Thus, in the lowest density stratum there is nearly a nine-fold increase in the proportion Black household numbers from primary to secondary stage (.011 to .090). In the high density stratum the increase is closer to a twofold one (.072 to .190). The proportion Black at the two stages unweighted for the disproportionate allocation is 3 percent and fifteen percent. Comparison of these figures with the estimates for the Epsem design,i.e. 2 percent and thirteen percent indicates the reduction in screening achieved by disproportionate allocation.

As in the pilot study the percentage Black households varies over the strata, although here the gains in distinguishing the medium and low density strata are more evident. Across the three strata ,the percentage of Black secondary numbers varies in an approximate 2:1.5:1 ratio. The three strata also differ in the total proportion of secondary numbers that are assigned to residences. The high density Black stratum has larger proportions of secondary numbers assigned to businesses, probably reflecting the urbanization levels.

Table 3 demonstrates that lower proportions of nonresidential numbers (.378) are found in the half-cluster (100 series) in which the primary number fell than in the other half-cluster (.409), but this difference (.03) is not statistically significant at the .05 level (standard error about .02). Similarly, the proportion of Black households is somewhat smaller in the adjacent 100 series (.125) than in the 100 series of the primary number (.133). Again, this difference is not likely to be found in most replications of the experiment. This is another perspective on the results in Table 2, showing only negligible reduction in the proportion eligible in 100 series adjacent to that of the primary.

Averages across clusters in the proportion eligible for the survey are not the only criterion of evaluation. In order to implement an epsem design within strata, each sample cluster in the design must have a sufficient number of Black households to support the designated number of sample Black households. Thus, the distribution over clusters of the proportion eligible is also of interest. Figure 1 presents the distribution over all the clusters of the proportion of Black households by stratum. The stability of these distributions varies because the number of sample clusters is about four times greater in the high density stratum than the other two (224 clusters to about 60). The shapes of the distributions, however, appear to be very different across the three groups. The low and medium density distributions are highly skewed, with sixty percent of clusters in the medium density stratum and sixty five percent of clusters in the low density stratum having 5 to 20 percent Black. These eligibility rates correspond to a maximum of 10 to 40 sample Black households for the 200 series clusters. Further, the low density stratum has several clusters that would not permit those cluster sizes (6 of the 63 clusters in that stratum are estimated to have fewer than 10 Black households). The distribution in the high density stratum is much more uniform (four of the 224 clusters are estimated to have less than 10 Black households).

These distributions deserve more discussion. Given our current understanding of the assignment of residential numbers to available banks of numbers, there is no reason to believe that within a wire center (or a prefix) that there are general tendencies to assign different residential areas to different 100 series. That is, within an exchange serving both Black and nonBlack households the hypothesis of assignment of numbers without regard to the race of the subscriber is a strong one. Stated alternatively, unless the exchanges are subdivided into wire centers that correspond to the residential locations of Black households, there is no a priori reason for large amounts of clustering of Black households within 200 series. Following this logic, the more uniform distribution in the high density stratum reflects, we believe, the variability in proportions Black among the telephone populations in the different exchanges in the stratum. If the prefix serves an area of high proportion Black households we would expect higher proportions of Black households, but we would expect the proportion Black households among all household numbers to vary only at random within selected clusters.

Firm evidence on this score can be assembled only through large samples within single exchanges, but if this reasoning is correct it implies that the advantage of the rejection rule based on Black households is absent in samples of single wire centers, and that the gains in the disproportionate sampling of high density Black clusters lies in populations consisting of several exchanges that vary in the proportion of Black households served.

5. Sampling Variance Properties

Even though it might be cost efficient to use large numbers of sample households per cluster (and a smaller number of clusters for a fixed sample size) and disproportionate allocation to strata, the overall precision of the sample is affected by the level of clustering effects on survey variables and the inflation of the variance due to weighting. *Ceteris paribus*, the larger the number of sample elements chosen per cluster the higher the design effect (the ratio of the sampling variance of the given design to that of a simple random sample with the same number of elements). The model often used is $Deff = 1 + \rho(b - 1)$, where Deff is the design effect, ρ is the intracluster correlation for the statistic, and b is the number of sample elements per cluster. Others have shown for many variables on the total U.S. household population that the intracluster correlations for the 100 series tend to be smaller than those generally found in area probability sample clusters (see Groves, 1978). This may not be the case for the Black population for 100 series and there are no empirical estimates available concerning intracluster correlations for 200 series clusters. The expectation prior to estimating sampling errors was that there would be no change in the intracluster correlations between the 100 and 200 series. This hypothesis reflects the understanding of assignment of telephone numbers within wire centers that was described above.

The average design effect for seven survey statistics is 1.28 for the 100 series and 1.30 for the 200 series. The 100 series average design effect as estimated from those cases which fell into the 100 series of the primary number while the cases from the entire 200 series were used for the average 200 series design effect. The average cluster size of completed interviews is 2.0 for the 100 series (coefficient of variation .043) and 3.4 for the 200 series (coefficient of variation, .029). These design effects reflect all the stratification, clustering, and weighting in the design and also the fact that the variability in the cluster sizes in the 100 series is greater because the rejection rule forced an equal number of sample Black households at the 200 series but not necessarily at the 100 series level. Given that the average design effects for the 100 series versus the 200 series are close to one another (1.28 to 1.30), the dominant influence on the sampling variance appears to be weighting, with little loss in precision due to cluster size alone (moving from the 100 to the 200 series clusters). However, some of this lack of movement is associated with the reduced coefficient of variation of the cluster size for the 200 series.

Table 4 presents the synthetic intracluster correlations by stratum for the same seven statistics used for the estimate of average design effects. These estimates are unweighted so as to remove the confounding effect of weighting on the synthetic estimates. The synthetic intracluster correlations were obtained from the design effect, following Kish's model of $Rho = (deff - 1)/(b - 1)$. The estimates in the table tend to be unstable due to the small number of clusters in each stratum, the small average cluster size of completed interviews, and its associated coefficient of variation. These sample design features complicate our inference about clustering effects in the 100 versus the 200 series. The 100 series estimates of intracluster correlation are somewhat higher than those in the 200 series. We believe that this reflects more a weakness in the synthetic correlation than a real difference in clustering effects. We believe that these estimates provide little evidence that there is a change in the intracluster correlation between the 100 and 200 series.

6. Optimal Design Features

The previous sections of the paper address the effect of alternative sample features on cost efficiency and sampling variance. Survey costs and errors are often combined at the design step to address whether "optimal" features of the survey can be identified. This approach attempts to identify the design which offers minimum variance for a fixed set of resources given to the survey. Given the data in this research we can estimate the optimal choices of two design attributes: a) number of sample elements per cluster, and b) allocation of the sample across strata.

To determine the optimal cluster size we use a total cost model, $C = C_0 + C_a a + C_b ab$, where C_0 is fixed costs, C_a is the sampling and screening costs for each sample cluster, of which a are selected, and C_b is the sampling, screening and interviewing costs associated with each interview obtained, of which there are b in each cluster. Because the proportion of black households vary across the three strata in the design, the C_a and C_b parameters vary across strata (see Table 5). The optimal cluster size is $\sqrt{(C_a (1-\rho))/(C_b \rho)}$, (Kish, 1965). Using those costs Table 5 presents estimated optimal cluster sizes for overall means and proportions with three alternative levels of intracluster correlation, .005, .01, and .02. (These values are similar to those obtained in the actual survey for attitudinal and behavioral variables). The C_a and C_b cost estimates for each stratum also appear. The table shows that the optimal cluster sizes are largest in the low density stratum, reflecting the high screening costs in that group. Note also that these cluster sizes are tend to be larger than those actually used in the survey, 6.5.

Note further that the optimal cluster sizes are similar for 100 and 200 series clusters, the loss of cost efficiency of the 200 series relative to that of the 100 is minor and similar optimal cluster sizes result. (The sampling variance estimates also imply that intracluster correlations in the 100 and 200 series clusters are similar).

The optimal cluster sizes in Table 5 generally exceed the levels that could be supported with 100 series cluster. That is, a large proportion of 100 series clusters would not have a sufficient number of black household numbers to fulfill the designated cluster size. This would require the use of weights in the estimation to reflect the higher probability of selection of numbers in such clusters, with attendant increases in sampling variance of the estimates. For that reason alone, the 200 series is favored. Even with 200 series, the specified sizes could not be obtained for some clusters in the low density stratum. (This suggests the optimal cluster size solution should be altered to reflect the capacities of the clusters).

The second design decision evaluated is the choice of allocation across strata. The survey used sampling fractions in the ratio of 3:2:1 from the high density to the low density stratum. We explore the optimal allocation across strata, assuming that the optimal cluster sizes were chosen in each stratum (as shown in Table 5). Given a fixed cluster size in each stratum, b_h, we set the sampling fraction in the h-th stratum, f_h, proportional to $\sqrt{(Deff_h S^2_h)/(C_{ha}/b_h)}$, where $Deff_h$ is the design effect for the statistic in the h-th stratum, S_h^2 is the element variance in the h-th stratum, C_{ha} is the sampling and screening costs for clusters in the h-th stratum, and b_h is the number of sample elements per cluster in the h-th stratum.

Table 6 presents optimal ratios of sampling fractions in the three strata for various combinations of element variances in the three strata and various ρ values. The table shows that the optimal allocations across strata are relatively insensitive to changes in ρ values (for the range of ρ values that are likely given this design). If the strata with higher densities of black persons have element variances at least equal to that of the low density stratum, an oversampling of those strata is desirable. (This reflects the much lower costs of those strata). The 3:2:1 ratio of sampling fractions is best for when the ratio of standard deviations is about 1.7:1.5:1. An examination of the data obtained from the survey suggests that many variables have ratios of standard deviations across the three strata close to 1:1:1. For such variables the optimal ratio of sampling fractions is 1.7:1.4:1, given the optimal cluster sizes shown in Table 5. (With the cluster size of 6.5 actually used in each stratum, the optimal fractions have the ratio 2.5:1.6:1). Both these ratios of sampling fractions suggest that the oversampling actually used created a loss of precision per unit cost, relative to that corresponding to the optimal fractions.

7. Summary

In the context of a two stage RDD design, this paper has demonstrated large gains in the use of Black household rejection rules, increasing the cluster size from 100 to 200, and disproportionate allocation across strata differing in density of the Black population. The PPS sampling technique of the Black household rejection rule acts to locate clusters of higher density of Black households, thus dramatically reducing screening costs. The residential segregation of the Black population enables the rejection rules to be effective. The use of clusters of 200 adjacent numbers (instead of the traditional 100) allows the researcher to select larger numbers of Black households from each cluster, without incurring large reductions in the proportion eligible among generated sample numbers. Given the results of the study optimal cluster size was shown to be even larger than used in the production study. The stratification used in the design identified areas that had eligibility rates as much as twice that expected in the full population. Optimal allocation to the strata was shown to be sensitive to element variances across the three strata; given the empirical results of the study, a 1.7:1.4:1 ratio of sampling fractions appeared desirable, given optimal cluster sizes.

References

Blair, J. and R. Czaja, Locating a Special Population Using Random Digit Dialing, Public Opinion Quarterly, Vol.46, 1982, pp. 585-590.

Groves, R.M., An Empirical Comparison of Two Telephone Sample Designs, Journal of Marketing Research, Vol.XV, 1978, pp. 622-31.

Kish, L., Survey Sampling, John Wiley and Sons, Inc., 1965.

Thornberry, O.T. and J.T.Massey, Coverage and Response in Random Digit Dialed National Surveys, Proceedings of Section on Survey Research Methods, American Statistical Association, 1983.

Waksberg, J., Sampling Methods for Random Digit Dialing, Journal of the American Statistical Association, Vol. 73, No.361, 1978, pp. 40-46.

Table 1
Pilot Study
Disposition of Secondary Numbers Selected within 100,200,400 series by Stratum

Stratum and Disposition	Proportion of All Numbers Selected*		
	100 Series	200 Series	400 Series
High Density Black Stratum			
Black Households	.205	.201	.214
Don't Know Race	.028	.029	.032
Nonblack Households	.316	.279	.275
Nonresidential/ Nonworking	.451	.491	.479
Number of Cases	(395)	(806)	(1163)
Medium Density Black Stratum			
Black Households	.104	.080	.076
Don't Know Race	.030	.018	.020
Nonblack Households	.494	.443	.420
Nonresidential/ Nonworking	.372	.459	.484
Number of Cases	(231)	(560)	(878)
Low Density Black Stratum			
Black Households	.085	.084	.069
Don't Know Race	.014	.028	.027
Nonblack Household	.532	.577	.607
Nonresidential/ Nonworking	.369	.311	.297
Number of Cases	(141)	(286)	(491)
Total			
Black Households	.134	.124	.115
Don't Know Race	.024	.025	.026
Nonblack Households	.442	.431	.448
Nonresidential/ Nonworking	.400	.420	.411
Number of Cases	(767)	(1652)	(2532)

*Weighted Estimates

Table 2
Production Study
Disposition of Numbers Selected by Stratum, and Primary/Secondary Number Status

Stratum and Disposition	Primaries	Secondaries
High Density Stratum		
Black Households	.072	.190
Don't Know Race	.035	.027
Nonblack Household	.219	.352
Nonresidential/ Nonworking	.674	.431
Number of Cases	(3,128)	(6,671)
Medium Density Stratum		
Black Households	.032	.141
Don't Know Race	.020	.018
Nonblack Household	.188	.469
Nonresidential/ Nonworking	.760	.372
Number of Cases	(1879)	(2,375)
Low Density Stratum		
Black Households	.011	.090
Don't Know Race	.019	.023
Nonblack Household	.199	.505
Nonresidential/ Nonworking	.771	.382
Number of Cases	(6,116)	(3,987)
Total for Epsem Design*		
Black Households	.021	.129
Don't Know Race	.021	.023
Nonblack Household	.200	.454
Nonresidential/ Nonworking	.758	.394
Proportion Black Households for Disproportionate Design	.031	.150
Number of Cases	(11,123)	(13,033)

*Weighted Estimates

Table 3
Production Study
Dispsition of Secondary Numbers by Whether in
Same 100 Series as Primary Numbers

Status	Disposition	
	Same 100 series as Primary	Adjacent 100 series
Black Households	.133	.125
Don't Know Race	.024	.022
Nonblack Household	.465	.444
Nonresidential/ Nonworking	.378	.409
Number of Cases	(6,522)	(6,511)

Table 4
Production Study
Synthetic Intracluster Correlations
for 100 and 200 Series Clusters for Seven Statistics by Stratum

Statistic	Synthetic Intracluster Correlation*					
	High Density Black Stratum		Medium Density Black Stratum		Low Density Black Stratum	
	100 Series	200 Series	100 Series	200 Series	100 Series	200 Series
Proportion Very Satisfied With Life As A Whole	.021	-.002	-.172	-.042	-.238	-.116
Proportion Who Think They Are Better Off Financially Than One Year Ago	.113	.075	.094	.069	.206	.049
Proportion Who Will Vote For Mondale	.189	.021	.086	-.087	-.436	-.046
Proportion Who Attend Church	.013	.017	-.009	-.078	.035	-.110
Proportion In Same City Or Town All Of Life	-.078	.001	.058	.114	.221	.248
Proportion Voted In 1980 Presidential Election	-.045	-.035	-.101	-.013	.364	.356
Proportion Who Think Reagan Will Be Elected President	-.045	-.045	-.545	-.078	.124	-.105
Average	.024	.005	-.084	-.016	.039	.039

*These estimates are unweighted.

Table 5
Cost Parameters and Optimal Number of Sample Elements Per Cluster, By
Stratum for 100 and 200 Series Clusters and Different ρ Values

Stratum and Cluster Definition	ρ Value			Cost Parameters	
	.005	.01	.02	c_{ha}	c_{hb}
High Density Stratum					
100	15.9	11.2	7.9	$50.81	$40.11
200	15.9	11.2	7.9		$39.78
Medium Density Stratum					
100	22.4	15.8	11.1	$114.09	$45.18
200	21.3	15.0	10.6		$50.00
Low Density Stratum					
100	29.8	21.0	14.8	$309.98	$69.52
200	29.9	21.1	14.8		$69.18

Table 6
Optimal Allocation of the Sample Across Strata for Overall Means, Given Optimal Cluster Sizes in Each Stratum, For Various Relative Standard Deviations Across Strata and Values of Intracluster Correlations

Ratios of Within Stratum Standard Deviations (High:Med:Low)	Ratios of Optimal Sampling Fractions (High:Med:Low)
	ρ = .005
3 : 2 : 1	5.2 : 2.7 : 1
1.7 : 1.5 : 1	3 : 2 : 1
1 : 1 : 1	1.7 : 1.4 : 1
.33 : .5 : 1	.6 : .9 : 1
	ρ = .01
3 : 2 : 1	5.2 : 2.7 : 1
1.7 : 1.5 : 1	3 : 2 : 1
1 : 1 : 1	1.7 : 1.4 : 1
.33 : .5 : 1	.6 : .9 : 1
	ρ = .02
3 : 2 : 1	5.1 : 2.7 : 1
1.8 : 1.5 : 1	3 : 2 : 1
1 : 1 : 1	1.7 : 1.3 : 1
.33 : .5 : 1	.6 : .9 : 1

FIGURE 1
PRODUCTION STUDY
PROPORTION OF BLACK HOUSEHOLDS WITHIN 200 SERIES CLUSTERS