BUREAU OF THE CENSUS

STATISTICAL RESEARCH DIVISION REPORT SERIES

SRD Research Report Number: CENSUS/SRD/RR-85/02

THE FEASIBILITY OF INTERVIEWER VARIANCE ESTIMATION IN THE CPS: FINAL REPORT

by

Paul Biemer
Statistical Research Division

John Bushery
Donna Kostanich
Fay Nash
Statistical Methods Division

Bureau of the Census
Room 3573, F.O.B. #3
Washington, D.C. 20233

(301)763-5350

Recommended by: Kirk M. Wolter

Report completed: March 13, 1985

Report issued: April 23, 1985

The Feasibility of Interviewer Variance Estimation in the CPS: Final Report

by

Paul Biemer, John Bushery,

Donna Kostanich and Fay Nash

February 28, 1985

Table of Contents

	Page
I. Introduction	
A. General	1
B. Assumptions Made	2
II. Summary of Findings and Recommendations	4
A. Summary of Findings	4
B. Recommendations	6
III. General Design for Interviewer Variance Estimation	9
A. Design Considerations	9
B. The Interpenetration Scheme	11
C. Precision of the Estimates	16
IV. Method for Interpenetrating Interviewer Assignments	18
A. Description	18
B. Procedure for Pairing Interviewers	28
C. Sample Selection	29
D. Interpenetration of Assignments	30
E. Potential Problems and Remedies	31
F. Pretest of Interviewer Pairing Procedures	34
V. Field Implementation	35
A. Notification to the Regional Offices as to the Sample	35
B. Procedure for Interpenetration of Assignments for Selected Interviewer Pairs	35
C. Quality Control on Interpenetration Operation	35
D. Controlling Subsequent Assignments of Segments Initially Selected for Interpenetration	36
VI. Feasibility of the Interpenetration Scheme	39
A. Interpenetrable Areas	39
B. Impact of Redesign Changes	42
C. Travel Increases	42
D. Overall Travel Increases for ℓ Interpenetrated Interviewer Pairs	47
E. Interviewer Costs	48

Table of Contents (Cont'd.)

Appendices

		Page
Appendix I	Derivation of Segment-to-Segment Travel Components for Interpenetrated Areas......................	52
Appendix II	Derivation of Home-to-Segment Travel Components for Interpenetrated Areas......................	57
Appendix III	Computations of Travel Times......................	62
Appendix IV	Determination of Travel Time Components Given ℓ Pairs Selected for Year i......................	66
Appendix V	Yearly estimates of Interviewer Variance..........	71

I. Introduction

A. General

This report summarizes the preliminary results of a study to investigate the feasibility of estimating interviewer variance in the CPS. One option of this program would be to provide estimates of interviewer effects periodically on an ongoing basis and would be designed as a permanent part of the CPS. Another option would be a one-time study to estimate the interviewer correlated components associated with the labor force and other characteristics.

A major goal of this study is to determine what design changes, if any, are required for CPS in order that a program of interviewer variance estimation be feasible. These design changes would then be considered for future CPS redesigns. Some questions for this study are:

1. Is a one PSU per stratum design more or less compatible with a system of interpenetrated interviewer assignments than a two PSU per stratum design?

2. For what proportion of the CPS workload are interviewer assignments so widely dispersed as to make interpenetration impractical?

3. What are the increases in travel and other administrative expenses that result from interpenetrating workloads? What is the expected cost of an ongoing program and a one-time study?

4. Can a system of interpenetrated interviewer assignments be compatible with the rotation/panel sample design of CPS?

In order to determine the feasibility of an interviewer variance estimation program, it was necessary to develop a general design for a system of interpenetrated interviewer work assignments which could feasibly be used for the program. The proposed system is fully described in Sections III and IV.

Section II summarizes the findings of this investigation and gives the group's recommendations for CPS.

Also in Section III, the proposed scheme for interviewer assignment interpenetration is described, accompanied by a method for phasing in the scheme. The form of the estimator of interviewer variance and the precision of the estimate are also discussed in this section.

In Section IV the procedure for pairing and sampling neighboring CPS interviewer assignments for interpenetration is discussed, as are the problems of interviewer attrition, attrition of units, temporary loss of interviewers, etc.

Section V describes the field implementation aspects of the design: procedures for interpenetrating assignments, quality control, and administrative problems.

Finally, in Section VI, the feasibility of interpenetration in CPS PSUs is examined for all areas of the country and travel costs are estimated.

B. Assumptions Made

The majority of the work for this investigation was completed by the Interviewer Variance Workgroup in 1982. That year, a preliminary version of this report was circulated to the Current Survey Redesign Steering

Committee for comments. That version was complete except for a discussion of the precision of the estimator of interviewer variance. The Interviewer Variance Workgroup recommended that the report not be finalized until formulas for the variance of the proposed interviewer variance estimator could be derived and a discussion of the precision of the estimators and the sample size requirements could be incorporated into the report. The basic methodology for predicting the precision of the interviewer variance estimator was recently completed and is documented in [2]. In Appendix V, this methodology was applied to the interviewer variance estimator proposed for CPS. Thus, this final report now contains a discussion of the sample size requirements and costs.

In 1982, major decisions regarding the sample design for the CPS redesign had not yet been made. Therefore, certain assumptions were made about the current design so that progress could be made on this report.

First, it was assumed that the basic methodology would remain unchanged; that is, that there would be no substantive differences from the data collection methods used since the 1970 redesign. This includes the basic sample rotation plan and the PSU replacement plan.

The redesigned CPS sample was expected to be similar to the A/C/D design, the exception being that the reliability requirements for states would be slightly stricter (8% CV on the estimated annual average of the number of unemployed as compared to the previous 10% CV). Therefore, the A/C/D sample was used as a test vehicle for the feasibility of the interviewer variance program for the redesigned CPS sample.

Since this report is concerned primarily with whether a program of interviewer variance estimation is feasible, many of the details of the various phases of the program are not discussed once the feasibility of each phase is established. Furthermore, no claim is made that the approach taken in this report is optimum although we believe it is the "best" approach among all the alternatives considered.

II. Summary of Findings and Recommendations

A. Summary of Findings

To briefly summarize the findings of this investigation, the questions stated in Section I are now discussed.

1. The feasibility of the interviewer variance estimation program should not affect a decision to sample two PSUs per stratum rather than one PSU per stratum in the redesigned CPS. Even with a two PSU per stratum design, it is likely that each NSR PSU would be a single-interviewer PSU, the same as for a one PSU per stratum design. In either case, the only interpenetrable NSR PSUs would be those that are contiguous to another sample PSU. As estimated from the A/C/D design, non-contigious NSR PSUs constitute less than 20 percent of the interpenetrable sample and consequently any differences between the two designs (i.e., one or two PSUs per stratum) would be relatively small.

2. An estimate of interviewer variance based upon the interpenetrable areas in the A/C/D CPS design would represent about 91%

of the correlated component of variance contributed by all interviewers. Futhermore, interpenetrable areas comprised about 87% of the CPS sample. This is enough to allow the estimated interviewer variance component to be used as a reasonable indicator of the CPS data quality.

3. A system of interpenetrable assignments was designed that could feasibly become a permanent part of the CPS. Few changes are required in the field operations. The travel for a Pattern I interviewer (an interviewer whose assignment is interpenetrated) would increase in about one-fourth of his/her assignment. (Recall that about 70% of the interviews in the CPS are conducted by telephone). Special procedures would have to be developed for assigning segments of households to the interviewers in accordance with the study design.

4. Based upon the theoretical work reported in [2] and Appendix V, it is estimated that approximately 100 pairs of interviewer assignments must be interpenetrated in order to obtain annual estimates of interviewer effects on labor force status variables with acceptable levels of precision (i.e., a 20% coefficient of variation on the intra-interviewer correlation coefficient). This is based upon a procedure which randomly assigns <u>segments</u> of households (not individual households) within interpenetrated assignment pairs. For an ongoing program of interviewer variance estimator, the increase in cost of the field work for a sample of this size is $118,000 per year. For a one-time study, the cost is $239,000 for a 27-month study producing two annual estimates of in-

terviewer variance. These costs do not include the costs of planning, preparation, and analyses.

The methodology developed to answer the question of the precision of the estimator makes a number of statistical assumptions, many of which could be tested in a small pilot study. In addition, a pilot study conducted in several regional offices could be useful to validate the model-based answers to questions regarding the operational feasibility and costs of the interpenetration and estimation procedures. The pilot study would also provide some much needed information on the magnitude of the interviewer variance components for CPS labor force variables. The detailed recommendations of the group are listed below.

B. Recommendations

The majority of the work for this project was done during the early stages of the current survey redesign. At that time, there was concern that certain sample designs under consideration would adversely impact our ability to estimate interviewer variance in the current surveys. The first recommendation of the group addresses that concern.

1. Proceed with the redesign without concern for enhancing the compatibility of an interviewer variance estimation program in the CPS. The feasibility of interviewer variance estimation is not constrained by any of the proposed CPS sample redesigns.

If, based upon the results of this report, the Bureau wishes to pursue an ongoing program or a one-time study to estimate the interviewer variance in the CPS, the assumptions made in this report related to costs and precision should be re-examined and validated.

This might be accomplished through a small-scale pilot study in one or more RO's in combination with a computer simulation study of the estimators and their variances. Additionally, the assumptions made about the redesign made in 1982 should be checked and verified in light of the actual redesign and, if necessary, the results of this feasibility study should be modified to account for the current design. Therefore, the group recommends that a task force be formed to address the following objectives:

a. test and validate (via a small pilot study) the cost assumptions and projections made in the report,

b. demonstrate in a pilot study the operational manageability and practicality of the procedures,

c. obtain data on the size of the interviewer variance component for CPS labor force variables, and

d. validate the computational variance formulas used to determine the recommended sample size for the report.

C. <u>On-going Program, One-time Study, or No Study?</u>

The work group intentionally did not address in our recommendations the question - "Should, the CPS interviewer variance component be estimated?" The answer depends largely upon transient factors such as budgetary restraints and research priorities. Further, given there is a Bureau commitment to investigate interviewer effects, the question as to whether an on-going program or a one-time study is preferred also cannot be easily answered. The answer depends upon the available resources as well as the intended use of the estimates as the following points should illustrate.

1. The best known and classic example of the utility of estimates of interviewer effects is the enumerator variance study of the 1950 census. In 1950 enumerator assignments were interpenetrated and the enumerator variance was analyzed. The analyses revealed enumerator effects so large that subsequent censuses (in the U.S. as well as in many other countries) have used self-enumeration as the primary method for census data collection.

   Thus interviewer variance studies can substantially influence decisions regarding modes of data collection for a survey. For CPS and other current surveys, this information could be useful in deciding whether the Bureau should adopt centralized telephone interviewing procedures, cold-contact random digit dialing methods, or computer assisted personal interviewing technologies.

2. The interviewer variance component can be a major term in the standard error of an estimator (such as the unemployment rate), and may contributes more to state estimator standard errors than to national estimator standard errors. It is widely recognized that for self-representing areas in the CPS sample, the current estimators of standard error do not reflect interviewer variability and consequently understate the true variability. The amount of this downward bias depends upon the magnitude of the interviewer component. We currently have no information on how serious this problem is. Studies of the interviewer component could address this concern.

3. Interviewer variance estimates can allow the monitoring of CPS interviewer performance. By tracking the interviewer component over the years, it should be possible to determine whether interviewer effects

are increasing (performance is deteriorating) or decreasing (performance is improving). It should, therefore, be possible to investigate what impact changes in interview instructions and procedures, supervisor-to-interviewer ratios, incentives for reducing interviewer attrition, etc. have on data quality.

The question of whether to choose an on-going program or one-time study depends upon which of these uses is targeted. For example, an on-going program would allow uses (1) - (3) as well as many others not mentioned. A one-time study may not be useful for objective (3) but would allow (1) and possibly (2).

III. General Design for Interviewer Variance Estimation

A. Design Considerations

In selecting a design for interviewer assignment interpenetration, several theoretical and practical requirements were identified. First, the requirements of the estimation procedure specify that all interviewer assignments be paired together in some fashion[1/], preferably within PSUs, and that a random sample of these pairs be interpenetrated. The practical requirements of the method dictate that these assignment pairs should be "nearest neighbors" in order to reduce the cost of travel for the interviewers.

A second requirement is that the "general conditions" for interpenetrated interviewer assignments should be essentially the same as

[1/] Although we considered only pairs, any number of assignments can be grouped together for interpenetration. It is shown in [2] that pairing assignments is optimal, however.

for noninterpenetrated assignments, with the exception, of course, of travel within assignments. This means that the assignment of households to the interviewers and the handling of the assignments should be the same administratively for both types of assignments. Furthermore, in order to further reduce the possible effects of conditioning of the interviewers, interviewers should be rotated into and out of the study so that the same sample of interviewer pairs is not always interpenetrated.

Finally, the interpenetration scheme must be compatible with the rotation-panel survey design of the CPS. For example, one interviewer should be allowed to conduct all eight interviews for a particular household; the effects of rotation group biases on the interviewer variance estimates should be minimized; problems such as new construction, demolitions, attrition, etc. must be efficiently handled; and so on.

Indeed, in some parts of the country (especially in the West), interpenetration of interviewer assignments is impractical due to the vast distances between neighboring assignments. Therefore, a central question to this investigation is "in what proportion of the interviewer assignments is interpenetration feasible?" Using very conservative assumptions, 91% was the estimated proportion. This is discussed in detail in Section VI.

Once satisfied that interpenetration was feasible for a sufficient proportion of the CPS workload, the committee addressed the question of "how can interpenetration of interviewer assignments be accomplished?" The next section describes one possible scheme for interpenetration.

B. The Interpenetration Scheme

The interpenetration scheme described below pertains only to those interviewer pairs for which interpenetration is practical. The resulting estimates will of course represent only the population of interpenetrable assignments.

An important feature of this interpenetration scheme is that either the household or the segment could be used as the unit of randomization for interviewer interpenetration. Because of the inherent administrative problems with interpenetration of households within segments, it was decided that the feasibility of interpenetrating segments of about four housing units would be investigated in this report.

Phase-In of the Scheme

Let 2L denote the total number of interviewer assignments in interpenetrable areas of the U.S. Suppose that the 2L assignments are paired together by some method such as that described later in Section IV. A random sample of ℓ of these L pairs is drawn (with equal probability) for interpenetration. The interviewers associated with these ℓ pairs will have their assignments interpenetrated and will be referred to as Pattern I interviewers (following the terminology in [1]) and those interviewers associated with the other L-ℓ pairs will maintain their usual assignments and will be referred to as Pattern II interviewers.

To phase-in the interpenetration scheme, only the incoming CPS rotation groups are interpenetrated. To accomplish this, incoming segments are initially allocated to the interviewers as usual. Then, those incoming segments in Pattern I assignments are randomly reassigned to the two interviewers of the pair. For example, let A and B identify the two interviewers of a Pattern I pair. Suppose each interviewer obtained two new segments from the incoming rotation group and denote these by A1, A2 and B1, B2, respectively. Then, the interpenetration procedure would reassign either A1 or A2 (decided by random selection) to B and either B1 or B2 to A.

This procedure of interpenetrating the incoming segments for the ℓ pairs of Pattern I interviewers continues for each month. The Pattern II interviewers have normal assignments. Of course, once an incoming segment has been reassigned to a Pattern I interviewer, that segment remains assigned to the interviewer for all eight CPS interviews.

For an ongoing program of interviewer variance estimation, a new sample of Pattern I interviewers is selected each year. After twelve months from the commencement of phase-in, the incoming rotation group assignments of the previous sample of Pattern I interviewer pairs are no longer interpenetrated. Now, the incoming rotation groups for the new sample of Pattern I interviewers are interpenetrated following the same procedure as described above.

The process of interpenetrating the incoming rotation group's segments continues for the second sample of interviewer pairs. Then, 12 months later a third sample of ℓ interviewer pairs is selected and so on. This introduction of new samples of Pattern I interviewers and phase-out of old interviewer samples is repeated after every 12 months of the survey.

Figure 1 summarizes this interpenetration scheme for the first 37 months of the program. The entries of the table are of the form a(b,c) where b denotes the CPS household sample, c denotes the rotation group number (c=1,...,8) and a denotes the sample of Pattern I interviewers or, equivalently, the year of the study. The symbol a(b,c) denotes the set of CPS households in sample rotation (b,c) assigned to the a-th sample of Pattern I interviewers.

Note from the table that after every 12 month cycle, a new sample of Pattern I interviewers is selected and the incoming rotation groups are no longer interpenetrated for the old Pattern I interviewers; however, previously assigned rotation groups will continue to be interpenetrated for the previous Pattern I interviewers. In fact, any Pattern I interviewers will have some part of their assignments interpenetrated for exactly 27 months.

FIGURE 1: Sample Rotation Chart for an On-going Program of Interviewer Variance Estimation for CPS

MONTH OF PHASE-IN	MONTH OF INTERVIEW	1	2	3	4	DORMANT PERIOD 1	2	3	4	5	6	7	8	5	6	7	8
1	January	1(1,1)															
2	February	1(1,2)	1(1,1)														
3	March	1(1,3)	1(1,2)	1(1,1)													
4	April	1(1,4)	1(1,3)	1(1,2)	1(1,1)												
5	May	1(1,5)	1(1,4)	1(1,3)	1(1,2)	1(1,1)											
6	June	1(1,6)	1(1,5)	1(1,4)	1(1,3)	1(1,2)	1(1,1)										
7	July	1(1,7)	1(1,6)	1(1,5)	1(1,4)	1(1,3)	1(1,2)	1(1,1)									
8	August	1(1,8)	1(1,7)	1(1,6)	1(1,5)	1(1,4)	1(1,3)	1(1,2)	1(1,1)								
9	September	1(2,1)	1(1,8)	1(1,7)	1(1,6)	1(1,5)	1(1,4)	1(1,3)	1(1,2)	1(1,1)							
10	October	1(2,2)	1(2,1)	1(1,8)	1(1,7)	1(1,6)	1(1,5)	1(1,4)	1(1,3)	1(1,2)	1(1,1)						
11	November	1(2,3)	1(2,2)	1(2,1)	1(1,8)	1(1,7)	1(1,6)	1(1,5)	1(1,4)	1(1,3)	1(1,2)	1(1,1)					
12	December	1(2,4)	1(2,3)	1(2,2)	1(2,1)	1(1,8)	1(1,7)	1(1,6)	1(1,5)	1(1,4)	1(1,3)	1(1,2)	1(1,1)				
13	January	2(2,5)	1(2,4)	1(2,3)	1(2,2)	1(2,1)	1(1,8)	1(1,7)	1(1,6)	1(1,5)	1(1,4)	1(1,3)	1(1,2)	1(1,1)			
14	February	2(2,6)	2(2,5)	2(2,4)	1(2,3)	1(2,2)	1(2,1)	1(1,8)	1(1,7)	1(1,6)	1(1,5)	1(1,4)	1(1,3)	1(1,2)	1(1,1)		
15	March	2(2,7)	2(2,6)	2(2,5)	1(2,4)	1(2,3)	1(2,2)	1(2,1)	1(1,8)	1(1,7)	1(1,6)	1(1,5)	1(1,4)	1(1,3)	1(1,2)	1(1,1)	
16	April	2(2,8)	2(2,7)	2(2,6)	2(2,5)	1(2,4)	1(2,3)	1(2,2)	1(2,1)	1(1,8)	1(1,7)	1(1,6)	1(1,5)	1(1,4)	1(1,3)	1(1,2)	1(1,1)
17	May	2(3,1)	2(2,8)	2(2,7)	2(2,6)	2(2,5)	1(2,4)	1(2,3)	1(2,2)	1(2,1)	1(1,8)	1(1,7)	1(1,6)	1(1,5)	1(1,4)	1(1,3)	1(1,2)
18	June	2(3,2)	2(3,1)	2(2,8)	2(2,7)	2(2,6)	2(2,5)	1(2,4)	1(2,3)	1(2,2)	1(2,1)	1(1,8)	1(1,7)	1(1,6)	1(1,5)	1(1,4)	1(1,3)
19	July	2(3,3)	2(3,2)	2(3,1)	2(2,8)	2(2,7)	2(2,6)	2(2,5)	1(2,4)	1(2,3)	1(2,2)	1(2,1)	1(1,8)	1(1,7)	1(1,6)	1(1,5)	1(1,4)
20	August	2(3,4)	2(3,3)	2(3,2)	2(3,1)	2(2,8)	2(2,7)	2(2,6)	2(2,5)	1(2,4)	1(2,3)	1(2,2)	1(2,1)	1(1,8)	1(1,7)	1(1,6)	1(1,5)
21	September	2(3,5)	2(3,4)	2(3,3)	2(3,2)	2(3,1)	2(2,8)	2(2,7)	2(2,6)	2(2,5)	1(2,4)	1(2,3)	1(2,2)	1(2,1)	1(1,8)	1(1,7)	1(1,6)
22	October	2(3,6)	2(3,5)	2(3,4)	2(3,3)	2(3,2)	2(3,1)	2(2,8)	2(2,7)	2(2,6)	2(2,5)	1(2,4)	1(2,3)	1(2,2)	1(2,1)	1(1,8)	1(1,7)
23	November	2(3,7)	2(3,6)	2(3,5)	2(3,4)	2(3,3)	2(3,2)	2(3,1)	2(2,8)	2(2,7)	2(2,6)	2(2,5)	1(2,4)	1(2,3)	1(2,2)	1(2,1)	1(1,8)
24	December	2(3,8)	2(3,7)	2(3,6)	2(3,5)	2(3,4)	2(3,3)	2(3,2)	2(3,1)	2(2,8)	2(2,7)	2(2,6)	2(2,5)	1(2,4)	1(2,3)	1(2,2)	1(2,1)
25	January	3(4,1)	2(3,8)	2(3,7)	2(3,6)	2(3,5)	2(3,4)	2(3,3)	2(3,2)	2(3,1)	2(2,8)	2(2,7)	2(2,6)	2(2,5)	1(2,4)	1(2,3)	1(2,2)
26	February	3(4,2)	3(4,1)	3(3,8)	2(3,7)	2(3,6)	2(3,5)	2(3,4)	2(3,3)	2(3,2)	2(3,1)	2(2,8)	2(2,7)	2(2,6)	2(2,5)	1(2,4)	1(2,3)
27	March	3(4,3)	3(4,2)	3(4,1)	2(3,8)	2(3,7)	2(3,6)	2(3,5)	2(3,4)	2(3,3)	2(3,2)	2(3,1)	2(2,8)	2(2,7)	2(2,6)	2(2,5)	1(2,4)
28	April	3(4,4)	3(4,3)	3(4,2)	3(4,1)	2(3,8)	2(3,7)	2(3,6)	2(3,5)	2(3,4)	2(3,3)	2(3,2)	2(3,1)	2(2,8)	2(2,7)	2(2,6)	2(2,5)
29	May	3(4,5)	3(4,4)	3(4,3)	3(4,2)	3(4,1)	2(3,8)	2(3,7)	2(3,6)	2(3,5)	2(3,4)	2(3,3)	2(3,2)	2(3,1)	2(2,8)	2(2,7)	2(2,6)
30	June	3(4,6)	3(4,5)	3(4,4)	3(4,3)	3(4,2)	3(4,1)	2(3,8)	2(3,7)	2(3,6)	2(3,5)	2(3,4)	2(3,3)	2(3,2)	2(3,1)	2(2,8)	2(2,7)
31	July	3(4,7)	3(4,6)	3(4,5)	3(4,4)	3(4,3)	3(4,2)	3(4,1)	2(3,8)	2(3,7)	2(3,6)	2(3,5)	2(3,4)	2(3,3)	2(3,2)	2(3,1)	2(2,8)
32	August	3(4,8)	3(4,7)	3(4,6)	3(4,5)	3(4,4)	3(4,3)	3(4,2)	3(4,1)	2(3,8)	2(3,7)	2(3,6)	2(3,5)	2(3,4)	2(3,3)	2(3,2)	2(3,1)
33	September	3(5,1)	3(4,8)	3(4,7)	3(4,6)	3(4,5)	3(4,4)	3(4,3)	3(4,2)	3(4,1)	2(3,8)	2(3,7)	2(3,6)	2(3,5)	2(3,4)	2(3,3)	2(3,2)
34	October	3(5,2)	3(5,1)	3(4,8)	3(4,7)	3(4,6)	3(4,5)	3(4,4)	3(4,3)	3(4,2)	3(4,1)	2(3,8)	2(3,7)	2(3,6)	2(3,5)	2(3,4)	2(3,3)
35	November	3(5,3)	3(5,2)	3(5,1)	3(4,8)	3(4,7)	3(4,6)	3(4,5)	3(4,4)	3(4,3)	3(4,2)	3(4,1)	2(3,8)	2(3,7)	2(3,6)	2(3,5)	2(3,4)
36	December	3(5,4)	3(5,3)	3(5,2)	3(5,1)	3(4,8)	3(4,7)	3(4,6)	3(4,5)	3(4,4)	3(4,3)	3(4,2)	3(4,1)	2(3,8)	2(3,7)	2(3,6)	2(3,5)
37	January	4(5,3)	3(5,4)	3(5,3)	3(5,2)	3(5,7)	3(4,8)	3(4,7)	3(4,6)	3(4,5)	3(4,4)	3(4,3)	3(4,2)	3(4,1)	2(3,8)	2(3,7)	2(3,6)

a(b,c) = set of CPS households for the ath sample of interviewers, the bth sample of hh's (set of 8 rotation groups), and the cth rotation group (c=1,...,8)

<u>Discussion of the Procedure</u>

The frequency that interviewers are rotated in and out of Pattern I is, to a large extent, arbitrary. However, the more often a "fresh" interviewer sample is introduced, the smaller is the proportion of an interviewer's assignment that is interpenetrated; but, at the same time, the greater is the complexity of the estimation and field procedures. Although the overall average increase in travel is not appreciably affected by rotation frequency, average travel for <u>individual</u> interviewers may be increased if a longer period of rotation is adopted.

This interviewer rotation - interpenetration scheme satisfies the theoretical and operational requirements stated earlier. In addition, there are a number of other desirable features of the design:

1. Assuming that an interviewer is selected for Pattern I no more than once every three years, at most only $\frac{1}{2}$ of his/her assignment will be interpenetrated, or, equivalently, only about $\frac{1}{4}$ of his/her assignment will be located outside his/her usual territory. Thus, the interviewer burden due to increased travel is reduced which also should reduce interviewer conditioning. On the average, about 44% of a Pattern I interviewer's assignment will be interpenetrated over the 27 months he/she is in a Pattern I sample.

2. Estimates of interviewer variance can be obtained by time-in-sample. This attempts to reduce the effects of rotation group biases on the estimates and may also allows comparisons to be made between panels.

3. Precision in the variance estimates "accumulates" over time. Thus, the cost and administrative burden of interpenetration is allowed to spread over a longer time period (for example, 27 months for a one-time study).

4. For an ongoing program, the number of interpenetrated assignment pairs can change from one year to the next. This could allow for adjustments in the precision of the estimates.

C. Precision of the Estimates

The form of the estimator of interviewer variance for the CPS under the proposed design is given in Appendix V. The precision of the estimator can be roughly approximated by the formulas provided in [2]. Appendix V extends the methodology to the proposed CPS interpenetration scheme. Table 12 in Appendix V was computed assuming entire interviewer assignments (i.e., 12 segments) would be interpenetrated. Table 1 below was computed using formulas which assume about 1/2 of each interviewer assignment (i.e., 6 segments) would be interpenetrated. This table would be appropriate for a one-time study, rather than a continuous sampling scheme as described below.

Table 1. Number of Pairs Needed to Achieve a 20% Coefficient of Variation on the Estimator of Interviewer Variance: One-time Study

Intra-interviewer correlation coefficient	Between Month-in-Sample Correlation Coefficient		
	.5	.6	.7
.001	263	289	314
.003	93	98	103
.005	72	75	77
.01	60	61	62
.02	55	55	56

Because accurate information is not readily available on the correlation between CPS labor force status recorded for the same household on different months in the sample, a range of values for this correlation is given in the table (columns). In addition, very little information is available on the size of the intra-interviewer correlation coefficient for labor force variables. Preliminary estimates from a recent random digit dialing study using the CPS questionnaire indicate that this coefficient can be as high as .01. However, other data from the 1970 census indicate much smaller coefficients - in the range of .001 to .005. Therefore, a range of values of this parameter is presented in the table (rows). A recommended number of interpenetrated interviewer pairs of 100, should be adequate for a wide range of the parameter values.

For the proposed ongoing program of interviewer variance estimation, a new sample of interviewers is chosen each year; however, once chosen, some portion of the interviewers' assignments would be interpenetrable for 27 months. Thus, for the second and subsequent years of the program, at least two (and occassionally three) samples of Pattern I interviewer can be "active" at once. Table 2 is completely analogous to Table 1 and gives the size of each annual sample of Pattern I interviewer assignment pairs for a 20% coefficient of variation of the estimator of intra-interviewer correlation coefficient.

Table 2: Number of Pairs Needed to Achieve a 20% Coefficient of Variation on the Estimator of Interviewer Variance: Ongoing Program

Intra-interviewer correlation coefficient	Between Month-in-Sample Correlation Coefficient		
	.5	.6	.7
.001	134	147	160
.003	47	50	52
.05	37	38	40
.01	31	32	32
.02	28	28	28

For this scheme, a sample size of 50 interviewer pairs per year should be sufficient for a wide range of parameters. This is roughly equivalent to the preceeding one-time study recommendation.

IV. Method for Interpenetrating Interviewer Assignments

A. Description

1. General Plan

As mentioned earlier, a sample of interviewers will be paired and within pairs assignments will be interpenetrated. Theoretically an interviewer could be paired with any other interviewer in the country. From a practical standpoint, pairing will be performed only within PSUs or groups of contiguous PSUs.

In order to minimize the additional travel costs associated with this interpenetration scheme, we plan to pair interviewers so that the sum of between-interviewer distances within pairs is minimized. Other constraints will probably be applied as well. It appears best to pair all eligible interviewers in the PSU, then select a sample of pairs to be used in the interpenetration study. We would interpenetrate assignments only for the sample pairs.

The actual interpenetration would require combining the incoming assignments of the two interviewers in the pair and systematically reassigning the segments between the two interviewers. Thus, each of the two interviewers will be assigned about half the combined assignment, with half of each interviewer's interpenetrated workload coming from his/her partner's assignment.

2. Possible constraints

Several constraints should be considered, and methods for applying them devised. Suggestions for these constraints are listed below.

a. Maximize Within-PSU Pairing

It is desirable, to the extent possible, to pair interviewers within the same PSU and to avoid pairing interviewers in different PSUs. There are two major reasons for this: (a) pairing interviewers in different PSUs generally is more expensive than pairing interviewers within the same PSU, due to added travel (b) between PSU pairing will add administrative complexities in the field procedure and (c) pairing interviewers across PSUs will cause problems if one of the PSUs is a rotating PSU. (Rotating PSUs are so small they cannot support the CPS and other current surveys for a full decade and must be replaced by another PSU when the available sample is exhausted. Because between-PSU pairing will be performed using adjacent PSUs, a rotating PSU can cause serious problems in costs and logistics if its replacement is not also adjacent to the opposing PSU in the pair.)

In view of the administrative complications, it is also not desirable to pair interviewers who work in different regions. This would occur only with pairings carried out between PSUs and would be unlikely to happen in any case. If RO personnel carry out the actual pairing operation, no between-RO pairings would be made. (Each RO would perform pairings using only interviewers within the region).

b. Minimize Between-Interviewer Distance

The most important constraint is probably that of minimizing between-interviewer distances within pairs. Because the regional office (RO) tries to make up assignments so that segments are located close to the interviewer's home (to minimize home-to-segment distance), pairing interviewers to minimize between-interviewer distance will tend to minimize the extra travel and costs required for the interpenetration study.

c. Equal Workloads

Another important constraint might be workload size. Pairing should be performed only between interviewers having about the same size workload. To achieve this, interviewers might be stratified by workload size before the pairing process, or interviewers with odd-sized workloads (extremely small or extremely large) might be ruled ineligible for the study. The former seems the preferable course of action to take. Fortunately odd-sized workloads may not be much of a problem since the ROs generally attempt to assign all interviewers the same size workload. Moreover, the interpenetration will be phased in gradually, one rotation group at a time.

When an interviewer is selected for the variance study he/she will participate in the study for a total of 27 months. Because a new sample of Pattern I interviewers will be selected each year (assuming an ongoing program), it is possible that the same interviewer may be selected in one sample during one year and another sample the subsequent year. For simplicity we shall assume the interviewer is selected in only one sample to participate in the study. In this simplified situation, one-eighth of the interviewer's assignments will be interpenetrated during the first month of participation, with this fraction gradually increasing to a maximum of one-half. Thus, no more than one fourth of a Pattern I interviewer's assignment will come from the opposing interviewer's area.

Table 3 illustrates the effect of pairing interviewers with different sized assignments. It illustrates how workloads change for a hypothetical situation in which interviewer A usually has a workload of 88 units and interviewer B usually has a workload of only 24 units. In this example interviewer A's workload drops by as much as 18% (16 units) while interviewer B's workload increases by 66% (the same 16 units).

Table 3.

Change In Workloads
Due to Interpenetration

Number of Months Participation in Variance Study	Proportion of Assignment Interpenetrated	Proportion of Assignment from Opposing Interviewer's Area	New Workload A	New Workload B
0	none	none	88 1/	24 1/
1	1/8	1/16	84	28
2	1/4	1/8	80	32
3	3/8	3/16	76	36
4 through 24	1/2	1/4	72	40
25	3/8	3/16	76	36
26	1/4	1/8	80	32
27	1/8	1/16	84	28
28	none	none	88 1/	24 1/

1/ Normal workload, without interpenetration.

d. Pay Level

The RO's pay interviewers at two levels, the regular rate and at a special, higher rate for interviewers assigned to hard-to-enumerate areas.

If a regular-rate interviewer were paired with a special-rate interviewer this could cause problems of morale, because the regular-rate interviewer would be asssigned some segments in the difficult-to-enumerate area but probably without receiving the special pay. The special-rate enumerator, on the other hand, would have a reduced workload in the difficult area, but probably would continue to be paid at the special rate. Even if the RO could work out an arrangement satisfactory to both interviewers, there might be administrative problems. One solution would be to stratify interviewers by pay rate (regular/special) before pairing, or if this is impractical, ruling special pay rate interviewers ineligible for the interpenetration study. This may be only a slight problem in practice. Table 3 indicates that only 4 of the 12 ROs currently have any special-rate interviewers on the payroll. New York and Detroit employ 10 each, Charlotte has 6, and Seattle has 23. Apparently all 23 special-rate interviewers in the Seattle RO are employed in Alaska, where all interviewers are paid the special rate. Thus the suggested stratification by pay rate would be automatically achieved. In the other ROs some additional care may be required in pairing interviewers. However, the difficult areas are usually central cities, which also may automatically provide the suggested stratification.

e. Other Constraints

There may well be other constraints which we should consider when pairing interviewers. One way to determine these would be to provide a description of the proposed study to the ROs and to Field Division and request any additional contraints they feel necessary. A special pilot study would uncover any problems or constraints in pairing interviewers. An additional constraint mentioned by Field Division is the composition of the area covered by the interviewer (with respect to demographic, economic, and/or sociologic characteristics). For example, some neighborhoods are composed primarily of white residents who might be hostile to a black interviewer. Similarly, some primarily black neighborhoods might not tolerate a white interviewer. In such circumstances pairing a black and white interviewer and interpenetrating their assignments could lead to increased difficulty for the interviewers in completing their assignments, increased type A rates and so on. The ROs would be aware of other such problems which we might fail to anticipate, so their participation in determining constraints for interviewer pairing would be useful. If a decision is made to carry out the study, Field Division should be contacted as soon as possible on this matter.

Table 4.

STAFF TOTALS FOR THE CPS PROGRAM AS OF JAN. 1,1981 ARE FOLLOWS:

OFFICE	INT.	SFR'S	SPARES	SPEC TECH	INT. SPEC*
BOSTON	160	14	0	0	0
NEW YORK	70	10	0	0	10
PHILADEPHIA	158	22	11	0	0
DETROIT	143	14	0	0	10
CHICAGO	112	21	1	1	0
KANSAS CITY	201	26	9	0	0
SEATTLE	175	20	4	0	23
CHARLOTTE	155	20	8	0	6
ALANTA	136	16	8	0	0
DALLAS	203	27	12	0	0
DENVER	260	23	28	2	0
LOS ANGELES	141	10	3	0	0
NAT.TOTALS	1923	231	84	3	49

* These are the interviewers in the higher pay category.

3. Development of a Pairing Algorithm

The algorithm used for pairing interviewers must use at least between-interviewer distance, and perhaps some or all of the other constraints discussed in IV.A.2. above. An algorithm based on distance alone could be developed fairly readily. In fact, Geography Division apparently has developed a computer program which will accomplish this with little or no modification. The complexity of the program would of course be increased by the addition of any other constraints.

The number of different interviewer pairs which can be formed from n eligible interviewers is $k = \frac{n(n-1)}{2}$. Because each PSU generally has a small number of interviewers, the number of possible interviewer pairs usually would be modest. However, some PSUs have 20 or more interviewers, which increases the possible number of pairs considerably. Table 5 shows the number of pairs which can be formed from a specified number of interviewers, up to 20.

From these k interviewer pairs we must select a single set of n/2 pairs (or (n-1)/2 pairs, if n is odd).

Table 5.

Number of Pairs which Can Be Formed from a Specified Number of Interviewers

Number of Interviewers	Number of Possible Pairs
2	1
3	3
4	6
5	10
6	15
7	21
8	28
9	36
10	45
11	55
12	66
13	78
14	91
15	105
16	120
17	136
18	153
19	171
20	190

If this is done by computer, all possible interviewer pairs can be combined and the set having the smallest total of between-interviewer distances selected.

This discussion has neglected the constraints other than distance, in part because not all these constraints have been identified. The discussion on the procedure for pairing interviewers, given below, will deal with these other constraints.

B. Procedure for Pairing Interviewers

1. Initial Pairing

The first step in pairing interviewers would be to obtain a list of all eligible interviewers from Field Division. The list must include for each interviewer: interviewer code, home address, and whatever information is required to apply any non-distance constraints. If any stratification is required, this should be performed before pairing. At the same time geographic coordinates should be assigned to each interviewer. This could be done by geo-coding the interviewer's address to the ED level, then matching to a file produced by Geography Division which has geographic coordinates for the centroid of each ED. These coordinates would be used in the pairing algorithm. Output would consist of a listing showing the interviewers paired, the coordinates of each interviewer in each pair and the between-interviewer distance for each pair. For a limited number of PSUs these coordinates might be graphed to provide a more easily comprehensible model of the pairing process, and as a means of checking the pairing. Other useful output would be a listing of all possible interviewer pairs showing the coordinates of each interviewer and the between-interviewer distance for each pair.

2. Entering Other Constraints

The other constraints might be applied after the initial pairing (based on distance alone) if they could not be incorporated in the pairing algorithm. A particularly attractive way to do this would be to transmit the final pairing within each PSU to the appropriate

RO for review and suggested revisions. The RO would indicate pairings which were unworkable, for whatever reason, and suggest revisions as appropriate. The revisions and reasons for making them would then be returned to headquarters, where the final decision on the pairing would be made. The ROs suggested revisions would serve as input in this decision-making process. This procedure makes use of the ROs special knowledge of conditions within the region. Not only can problems unforeseen by headquarters be anticipated, but the most efficient alternative pairings will be obtained. It may be even more effective to have each RO perform the initial pairing, although this may result in some loss of efficiency in minimizing between-interviewer distances.

C. Sample Selection

After a universe of interviewer pairs has been created, we can systematically select a sample of pairs to obtain the desired Pattern I sample size and notify Field Division of which interviewers are in sample.

For the ongoing program, we plan to select a new sample of interviewer pairs each year. This would tend to limit the amount of time an individual interviewer would be involved in the evaluation. It is possible that interviewers could be re-paired every year before the sample selection, with the new sample selected independently of all preceding samples. A serious disadvantage is that the same interviewer could be selected for Pattern I in two consecutive years, with a different partner in each year. With two different partners, the area an interviewer is responsible for could be triple that required without the variance study. This would increase interviewing costs for travel and put an extra burden on the interviewer. It has an advantage, in that a periodic revision of the pairing tends to correct for turnover of interviewers. A compromise procedure might be to revise the pairing for all interviewers not currently in Pattern I. Pairs of interviewers could then be selected using a procedure which would minimize overlap; i.e., interviewers currently in Pattern I would be less likely to be selected for the succeeding sample of the program.

D. Interpenetration of Assignments

We plan to instruct the RO to perform the actual interpenetration for Pattern I interviewers, using the procedure outlined below.

1. List each sample interviewer's incoming assignment, for one interviewer in the pair at a time. This is to be done each month, with the sample number, and each segment listed.

2. Take a random start (1 or 2) and assign the first sample segment listed to the first or second interviewer, consistent with the value of the random start (e.g., if random start = 2, assign the first segment to the second interviewer).

3. Continue to assign segments, as they are listed, alternating between the first and second interviewer. Thus if the random start were 2, the assignment of segments would be as illustrated in Figure 2.

Figure 2. Interpenetration and Assignment of Sample Segments

Segment Listed	Sample Number	Segment Number	Interviewer Assigned	Interviewer Code
1st			2	
2nd			1	
3rd			2	
4th			1	
5th			2	
6th			1	
7th			2	
8th			1	

This procedure assures that each month the incoming sample to be interpenetrated is distributed uniformly between the two interviewers. That is, each interviewer will receive half of his/her assignment from his/her own area and half from the other interviewer's area.

E. Potential Problems and Remedies

Several problems in the implementation of this study may have adverse effects on the estimates of interviewer variance or on the cost of the variance study.

1. Attrition of Interviewers

Ideally the same two interviewers should remain paired for the full 27 months of participation in Pattern I.

There will, however, be some turnover among interviewers and care must be taken when replacing interviewers in Pattern I.

The simplest method of replacing interviewers in the variance program is to follow the normal procedure of simply assigning all the original interviewer's segments to the replacement.

Since Field tries to hire a replacement interviewer who lives in the same general area as his/her predecessor, this will also help to control the additional travel costs associated with interpenetrating assignments. It must be assured, however, that only one interviewer will serve as the replacement for the segments in the variance study. Further, this interviewer should not be assigned interpenetrated segments from any other interviewer involved in the study. Above all, the replacement interviewer must not be the opposing interviewer in the pair. In some cases the RO may have to temporarily assign a resigning interviewer's assignment to more than one replacement interviewer until a suitable replacement is found. The problems involved when interviewers are temporarily unavailable are discussed below.

2. Temporary Loss of Interviewers

When an interviewer goes on vacation or for some other reason is unable to carry out his/her assignment for a given month, a replacement interviewer must be used for that month. If the missing interviewer is in Pattern I, any interviewer other than the opposing interviewer may be assigned the interpenetrated part of the missing interviewer's assignment. However, all the interpenetrated segments should be assigned to the same temporary replacement. It would be most desirable if the temporary replacement were assigned no other interpenetrated segments, but not absolutely necessary, as long as he/she is not assigned interpenetrated segments from both interviewers in a pair.

It may be necessary to maintain this restriction over the life of the pair's participation in the program. That is, once an interviewer has temporarily replaced one interviewer in a pair, he/she is always ineligible to replace the opposing interviewer in the pair (at least for interpenetrated segments).

In small PSUs in the variance study (with only one or two interviewers), obtaining a temporary replacement may be more expensive because an interviewer from outside the PSU will be needed, and higher travel costs will be involved.

3. Interviewers Moving Within the PSU

   If an interviewer moves to a new address within the PSU, the distribution of assignments may become less efficient because the interviewer must retain the interpenetrated segments assigned earlier. This probably will be a minor problem, because the general practice is for each interviewer to retain assigned segments thoroughout their participation in the CPS. Therefore, the variance study should introduce no new problems of control and assignment for the ROs.

4. Units Lost from Sample

   Occassionally units are lost from the CPS sample due to demolition, fire, or for other reasons. We don't expect these occurrences to pose any problems with respect to the program.

5. Redefinition of Interviewer Pairs

   In Section III.C. we proposed re-forming interviewer pairs every year, before selecting the sample of pairs. As mentioned earlier, this will help reduce problems caused by interviewer attrition, because the pairing will be updated to account for changes in the interviewing staff.

Redefining interviewer pairs will, however, require time and effort in headquarters and probably in the ROs, as well. For this reason it may not be cost effective to redefine interviewer pairs on an annual basis. Some analysis of the costs involved is required before a final decision can be made on this matter.

6. Expansion of the CPS Sample

Since 1970 there have been three expansions of the CPS sample design (the D-sample, the B-sample, and the E-sample). There also has been a major sample reduction. If similar changes were to occur after the 1980 redesign, this could have an impact on the workloads of interviewers. New PSUs may be introduced into the sample providing additional areas which are eligible for the interviewer variance estimation program or PSUs could be dropped and the number of interpenetrated areas may be reduced. At this point we can't anticipate any effects of possible sample expansions or reductions on the interviewer variance program. We should remain aware that such effects are possible.

F. Pretest of Interviewer Pairing Procedures

It would be worthwhile to pretest the various alternative procedures for pairing interviewers. For example, we can pair interviewers using distance as the only criterion for a sample of PSUs and send the results to the ROs for comments. We can request the ROs to compare these results with their preferred interviewer pairing. Without some sort of pretest of the pairing procedure, it will be very difficult to develop an efficient method to perform the pairing.

V. Field Implementation

A. Notification of the Regional Offices as to the Sample

As each new sample of Pattern I interviewers is selected, the regional offices will be notified as to which interviewers must have their assignments interpenetrated for the coming year. Since the interpenetration scheme will be applied for each month's incoming rotation, the regional offices must be advised of the specific incoming sample/rotation for the particular month.

B. Procedure for Interpenetration of Assignments for Selected Interviewer Pairs

The scheme for interpenetrating interviewer assignments was described in Section III. of this report. The regional office must record each month, on a sample selection form, the incoming segments for both interviewers in the selected pair and randomly (using a systematic sampling procedure) select half of the segments for assignment to one interviewer with the balance being assigned to the other.

C. Quality Control on Interpenetration Operation

In order to verify that the interpenetration operation was performed properly in the regional offices, it is necessary to implement a quality control operation. Since the interpenetration scheme proposed is not particularly complicated, it would seem appropriate to implement the same type of procedure that is performed on the basic CPS sampling operation. That is, the clerical staff in Washington or Jeffersonville should do a 100% verfication of the sample selection forms used for interpenetrating the incoming rotation segments for selected assignment areas on a monthly basis.

D. Controlling Subsequent Assignments of Segments Initially Selected for Interpenetration

Once the initial interpenetration scheme has been determined monthly for each incoming rotation, some method must be devised so that the same assignments are made to each interviewer (or their substitutes) for the second and subsequent enumeration periods. For example, if segment 2011 has been assigned as a result of the interpenetration operation to interviewer #1 for the incoming sample A48 in January 1984, it should be assigned to the same interviewer in February - April 1984 as well as January - April, 1985. If this is not possible (due to interviewer illness or attrition), the interpenetrated cases would be assigned to the substitute or replacment interviewer, which may be any interviewer other than the opposing interviewer of the pair.

The recommended method for controlling monthly assignments of segments in the Pattern I sample is to use the sample selected forms prepared monthly for interpenetrating the incoming rotation segments of selected assignment areas. The regional offices would refer to the forms from the appropriate months and pull all segments recorded on those forms. For each assignment area, these segments would be grouped into interviewer assignments using the sample selection forms to determine to whom the segments were initially assigned. After the interviewer assignments have been made for the Pattern I sample, the remaining segments can be assigned to interviewers using the offices' normal procedure.

The office procedures for an ongoing program are a bit more complex than those for a one-time study due to the fact that two samples (in some months three samples) may be active at once. During the first year of the program control will be relatively easy because the scheme will be in effect with only one sample of Pattern I interviewers. For example, if the study began in January 1984, sample A48/rotation 2 would be the incoming sample rotation to be interpenetrated in January. In February, sample A48/rotation 3 would be incoming and sample A48/rotation 2 would be scheduled for its second enumeration. Therefore, the offices would refer to their January and February sample selection forms when making their February assignments. Beginning in April 1985, the interpenetration scheme will be fully operative and for each month of interview the offices will have to refer to 8 months of sample selection forms. For example, in April 1985, sampling forms needed are those which would show the interpenetration scheme for sample A48/rotations 2-5; sample A49/ rotations 6-8; and sample A50/rotation 1. Sample A48/rotations 2-5 forms would be for the first sample of Pattern I interviewers; the remaining forms, for the second sample.

In January 1986, the interpenetration scheme would be in effect with three samples of Pattern I interviewers. Consequently the offices would have to refer to various months of sampling forms prepared over three years. For example, for January 1986 enumeration, the offices would refer to the sample selection forms prepared in October-December

1984, January and October-December 1985, and January 1986. These procedures could be simplified by automating the assignment control process using a microcomputer in the RO. In order to direct the offices to the appropriate months of sample selection forms, it is recommended that this information be included in the standard monthly office memoranda sent from Washington.

1. Attrition of Interviewers

As discussed in Section IV. E of this report, it is desirable that the same two interviewers should remain paired for the full 27 months of interpenetration in Pattern I assignments. This will not be maintained, however, whenever an interviewer leaves the survey.

Whenever an interviewer resigns, the regional office must determine whether or not that interviewer had been a member of a Pattern I pair. If so, a replacement interviewer must be assigned those segments which were part of the interpenetrated sample of segments in the former interviewers assignment. Erroneously assigning these segments to the opposing interviewer of the interpenetrated pair of interviewers would result in the eliminating of these cases from the study and should therefore be avoided. Because of this constraint, the offices may lose some flexibility in hiring and developing assignments for replacement interviewers.

2. Temporary Reassignments

Occasionally, the regional offices must rearrange asssignments due to temporary loss of interviewers, illness, vacations, etc., as well as for various activities such as Type A noninterview follow-up. In this situation, the office can make the reassignment to any interviewer <u>except</u> the interviewer that has been paired with the interviewer whose assignment is being reassigned. Again, this may limit the flexibility that the office normally has in making temporary adjustments to assignments.

The regional office must always be aware of how staff or assignment changes will affect the interviewer variance program and ensure that the guidelines established for the program are followed. The results of the interviewer variance program will be contaminated to the extent that the regional offices do not recognize when assignment changes have an impact on the interpenetration sample.

The personnel changes described in points 1 and 2 above will also result in additional record-keeping activities for both the regional offices and headquarters. All changes in assignments of interpenetrated segments must be documented and this information transmitted to Washington. Some system must be devised for this purpose.

VI. Feasibility of the Interpenetration Scheme

A relevant question of concern in implementing the proposed scheme for estimating interviewer variance is: will there be enough sample areas that can be interpenetrated without huge increases in travel costs? For reasons previously mentioned the A/C/D sample design of the CPS will be used to determine the feasibility and the associated cost of this proposed scheme.

A. Interpenetrable Areas

It seems reasonable that all sample PSUs with at least two interviewers could be interpenetrated with a minimal increase in travel costs. These PSUs were identified using the October

1977 A/C/D eligible sample housing units by PSU and by assuming that a PSU with more than 75 sample units could be classified as a multiple interviewer PSU. Use of this procedure yielded the following results:

1. 151 or 52% of the 292 SR PSUs could be interpenetrated, and
2. 76 or 20% of the 376 NSR PSUs could be interpenetrated.

As it turns out, these estimates actually understate the number of PSUs that could be interpenetrated. For example, some sample interviewer PSUs adjoin other PSUs which could provide a second interviewer to form an interpenetratable assignment pair. Therefore, PSUs with less than 75 sample units were more thoroughly examined. If the PSU was contiguous to another PSU and the land area was not very large (this was a subjective determination) then the PSU was considered interpenetrable. This resulted in 111 SR and 160 NSR additional interpenetrable PSUs, thus yielding the following number of interpenetrable areas:

1. 262 or 90% of the 292 SR PSUs;
2. 236 or 63% of the 376 NSR PSUs; and
3. 498 or 75% of all 668 A/C/D PSUs.

More important though, for purposes of both the estimate and cost, is the amount of the sample which could be interpenetrated. Therefore, based on the above PSU classifications and the October 1977 A/C/D PSU sample sizes, the following gives the workload available for interpenetration:

1. 87% of the total 66,000 A/C/D workload;
2. 96% of the 42,000 workload in SR PSUs; and
3. 70% of the 24,000 workload in NSR PSUs.

The baseweights for these PSUs were then used to determine the percentage of the total population that would be represented by an estimate obtained from only these interpenetrable areas. This resulted in 97% of the total population being represented and 98% and 95% of the population in SR and NSR PSUs, respectively being represented.

An interviewer variance component estimated from the above interpenetrable areas would represent 91 percent of the total interviewer variance for the total population. This is based on what the obtained coefficient of σ_b^2 would be for the interpenetrable areas versus what it should be for all areas. The coefficient is determined by:

$$C = 50^2 \sum_i (BW_i)^2 \left(\frac{WKLD_i}{50}\right)$$

where;

BW_i is the baseweight for the i^{th} PSU and

$\frac{WKLD_i}{50}$ is the expected number of interviewers for the i^{th} PSU.

This coefficient C is given in formula (5.6) of [1].

Note that the above estimates are rough approximations of what we would expect, especially since the actual number of interviewers per PSU could vary considerably from the expected number. Other limitations on the above estimates are that they do not account for either pairing interviewers within the same region or with equal salaries. These estimates also do not account for how PSUs with an odd number of interviewers would be handled.

B. Impact of Redesign Changes

Although we cannot guess at estimates of interpenetrable areas for the redesigned CPS, we can speculate as to what impact design changes may have on the above estimates of interpenetrable areas. Limitations imposed on the design of the D-sample resulted in slightly more SR PSUs and slightly larger PSU workloads than necessary. These inefficiencies can be contributed to constraints imposed on the formation of state strata and changes in PSU status from NSR to SR, sometimes resulting in more sample than necessary for A/C PSUs. The direction of redesign research was to aim for single-interviewer NSR PSUs, regardless of how PSUs are defined. Since the redesigned CPS allows for optimization that was not possible in the creation of the expanded D-sample, this could mean slightly fewer interpenetrable PSUs.

C. Travel Increases

A relevant concern in implementing this interpenetration scheme is the additional costs involved, particularly those due to increased traveling. Based on the interpenetrable classification of A/C/D PSUs given above and on a model developed by Field Division, we were able to obtain a reasonable estimate of the expected increase in traveling.

The 'travel model' specified in [3] is given as:

$$T = (\lambda_1 s_1 - \lambda_2) d_1 r_1 + 2 \lambda_2 d_2 r_2 + (\lambda_1 s_1) d_3 r_3$$

where

T = total interviewer travel time

λ_1 = average number of visits per segment

s_1 = number of segments assigned

λ_2 = average numer of trips from home

d_1 = distance from segment to segment

r_1 = minutes per mile for between segment travel

d_2 = distance from home to segment

r_2 = minutes per mile for home to segment travel

$d_3 r_3$ = average time spent traveling within a segment

The total travel time T is the sum of the following three travel components, respectively:

1) segment-to-segment travel (T_{ss})
2) home-to-segment travel (T_{hs}) and
3) within-segment travel (T_{ws})

In [3] the average values for these parameters and for the three travel time components for five "travel strata" are provided. The national average travel time per interviewer given in [3] is appropriate for the A/C/D sample. The values for the travel time components are appropriate for non-interpenetrated PSUs, but were used to construct new values of T_{ss} and T_{hs} for interpenetrated PSUs. T_{ws}, the within-segment component, will not change if the interpenetration is not carried out within segments. Therefore, the travel components T_{ss} and T_{hs} for the noninterpenetrated PSUs were adjusted to reflect the increases due to sample interpenetration. This was done separately for multiple and single interviewer interpenetrable PSUs, thus creating ten additional travel strata (see Table 7 in Appendix I).

All A/C/D PSUs were then classified into one of these fifteen strata and the stratum workloads were obtained in order to weight the various travel components.

The "travel model" given above was then expanded to accommodate the components for interpenetrated areas. The following gives a description of the new subscripts needed.

n refers to noninterpenetrable areas,

i refers to interpenetrable areas, and

(m) and (s) denote multiple and single interviewer PSUs respectively.

For noninterpenetrable PSUs, the segment-to-segment travel time is the same as given in [2] which is:

$$T_{ssn} = (\lambda_1 s_1 - \lambda_2)\, d_1\, r_1$$

For interpenetrated areas this component is the same for both multiple and single interviewer areas and will be denoted as $T_{ssi(m)}$ and $T_{ssi(s)}$, respectively. The terms are defined to be:

$$T_{ssi(m)} = T_{ssi(s)} = (\sqrt{I} + \sqrt{1-I})\,(\lambda_1 s_1 - \lambda_2)\, d_1\, r_1$$

where I is the proportion of an interviewer's workload in another interviewer's area. The derivation of $T_{ssi(m)}$ and $T_{ssi(s)}$ can be found in appendix I.

The home-to-segment travel time component for noninterpenetrable PSUs is also taken to be that stated in [3] which is:

$$T_{hsn} = 2\lambda_2 d_2 r_2$$

For interpenetrated PSUs this component will be different for multiple interviewer and single interviewer PSUs. For multiple interviewer PSUs, this term can be written as:

$$T_{hsi(m)} = 2 d_2 r_2 \lambda_2 (1 + I)$$

and for single interviewer PSUs, we have:

$$T_{hsi(s)} = 2 d_2 r_2 \lambda_2 [1 + (\frac{d_4}{d_2} - 1) I]$$

where d_4 is the distance between centers of paired PSUs and d_2 is as defined above. The derivation of $T_{hsi(m)}$ and $T_{hsi(s)}$ can be found in appendix II.

In order to determine the increase in travel time that could be expected for each type of interpenetrated area (multiple or single interviewer PSUs), the value of I needs to be obtained. Note that I is the proportion of an interviewer's workload in another interviewer's area and will usually be about half of the proportion of an interviewer's workload which is interpenetrated. This proportion of interpenetrated sample depends on the scheme that will be used for interpenetrating the sample. Section IV of this report gives a description of the scheme that has been suggested for this study. Also included in Section IV, Table 3, are the proportions of an interviewer's assignment that are interpenetrated by month of participation in the study.

Consequently, based on the proposed scheme the average proportion of an interviewer's assignment that is interpenetrated for the 27 affected months is approximately .4444. This gives an average value of I to be .2222. The general form of the estimated travel time components are shown in Table 7. The computations of the expected total interviewer travel time for multiple and single interviewer PSUs, denoted by $\bar{T}_{i(m)}$ and $\bar{T}_{i(s)}$, respectively, and the corresponding average travel time for <u>all</u> interpenetrable areas, denoted by $\bar{T}_i$, for I = .2222 are contained in appendix III. The obtained values are:

$$\bar{T}_{i(m)} = 1160.3 \text{ minutes}$$

$$\bar{T}_{i(s)} = 1216.0 \text{ minutes}$$

$$\bar{T}_{i} = 1170.7 \text{ minutes and}$$

The expected travel time per interviewer when no areas are interpenetrated was estimated to be 911.85 minutes in [3]. From the data used in this report this value was estimated to be slightly higher (913.6) (see computation of $\bar{T}$ in the appendix).

The following gives the expected percent increases in travel time per interviewer by type of area:

Type of Area	Percent Increase in Travel
Single-Interviewer Areas	(1216.0/913.6 - 1) 100 = 33
Multiple Interviewer Areas	(1160.3/913.6 - 1) 100 = 27
All Interpenetrable	(1170.7/913.6 - 1) 100 = 28

D. Overall Travel Increases for ℓ Interpenetrated Interviewer Pairs

The travel increases produced thus far have assumed that the entire interpenetrable sample would be interpenetrated. This type of scheme is not being considered in this report and consequently the above travel increases are applicable only for estimating expected increases per interviewer. The proposed scheme, would select only a subset (ℓ) of the total possible interpenetrable interviewer pairs (L). For these ℓ pairs, all new incoming rotation groups for 12 months would be interpenetrated. Thus, some portion of the 2ℓ interviewer assignments would be interpenetrated for 27 months. Now assuming an interviewer workload is 50 housing units, the expected number of total possible interpenetrable pairs (L) is estimated to be:

$$L = \frac{W_i}{2(50)} = \frac{56,887}{100} = 568.87$$

where W_i is the total possible interpenetrable sample given in Table 7.

In order to obtain the increase in interviewing costs associated with this scheme, the average monthly travel for <u>all</u> interviewers must first be determined. The general form for estimating these travel times is contained in Appendix III, case 5 and is denoted by $\ddot{T}'|\ell$. Note that this would be the average for the 27 affected months.

The actual scheme discussed in this report proposes that interviewer variance be estimated as an ongoing process. It suggests selecting ℓ pairs of interviewers each year and interpenetrate all new incoming rotation groups that year for each of these ℓ pairs. From Table 3 in Section IV one can easily see that any of the ℓ pairs selected in a given year will have some proportion of their assignments interpenetrated for 27 months. Also, if the selection of the new set of ℓ pairs is done independently each year, it is possible in any given year that anywhere from ℓ to 3 x ℓ interviewer pairs will have some sample interpenetrated. Note that the 3 x ℓ pairs could only occur when ℓ is less than or equal to L/3 (or 190, using the current data).

Consequently, the number of pairs, ℓ , selected for Pattern I each year will affect the per interviewer travel increases. Thus, the joint probability distribution of an interviewer being selected for Pattern I in three concurrent years can be used to approximate travel costs when ℓ pairs are selected. An explanation of how this was achieved is given in appendix iv. Also Table 11 in the appendix gives an example of the calculations involved in obtaining the average travel time for all interviewers, denoted by $T|\ell$, when ℓ=400 pairs are selected.

E. Interviewer Costs

Now that the increase in travel time has been determined, this can be used to obtain the associated cost per interviewer. Interviewing costs are a function of not only time spent traveling, but also miles traveled and the time spent conducting an interview. Therefore, the cost per interviewer can be expressed as a sum of the following terms:

C_{int} = mileage cost + cost of travel time + cost of interviewing time

Thus, the cost in dollars can be written as:

$$C_{int} = \bar{h}\,\bar{m}\,(.225/\text{miles}) + (\bar{T}/60)\,(5.50/\text{hr.}) + \bar{h}\,(\overline{TI}/60)\,(5.50/\text{hr.})$$

where

$\bar{h}$ is the average sample size per interviewer and is assumed to be 50.

$\bar{m}$1/ is the average miles traveled per household and assumed to be 6 miles.

T is the average time (minutes) per interviewer spent traveling and is assumed to be 913.6 minutes as determined in this section.

$\overline{TI}$1/ is the average time (minutes) spent conducting an interview and is assumed to be 21.1 minutes.

This formula gives the average interviewer cost to be, C_{int} = \$247.95 when no sample is interpenetrated. To obtain the cost per-interviewer when ℓ pairs are selected, replace $\bar{T}$ with $T|\ell$ for an ongoing scheme or replace $\bar{T}$ with $\bar{T}'|\ell$ for a one-time study. Now assume that the increase in travel time is proportional to the increase in mileage time, $\bar{m}$, then $\bar{m}$ is replaced with $(\bar{m})\ (T|\ell)/\bar{T}$ or $(\bar{m})\ (\bar{T}'|\ell)/\bar{T}$ as appropriate. Note that the time spent conducting an interview will not change. If ℓ = 400 pairs were selected, $T\,|(\ell = 400)$ was calculated to be 1210.4 minutes. Consequently the average cost per interviewer can be obtained by:

1/ Values were provided by Mary Ellen Beach from the Methods Research Branch in Field Division.

$$C_{int} \mid (\ell = 400) = (50)(6)(1210.4/913.6)(.225) + (1210.4/60)(5.50) + (50)(21.1/60)(5.50)$$

$$C_{int} \mid (\ell = 400) = \$297.09$$

Thus, the percent increase in the cost per interviewer is:

$$(297.09/247.95 - 1) \times 100 = 20\%$$

Table 6 gives a brief summary of the effect of selecting ℓ pairs for both an ongoing and a one-time interviewer variance study.

Table 6. Summary of Travel Times and Cost per Interviewer

Numbers of pairs selected ℓ	400	300	200	100	50	0
ONGOING STUDY						
Average Monthly Travel Time $\bar{T} \mid \ell$ (minutes)	1210.4	1149.5	1083.5	1039.9	958.75	913.6
Percent Increase in Travel Associated Cost	32%	26%	19%	10%	5%	0%
Average Total cost per Interviewer Per Month	$297.09	$287.01	$276.08	$262.90	$255.42	$247.95
Percent Increase in Total Cost	20%	16%	11%	6%	3%	0%
Increase in Total Cost Per Year 1/	$775,000	$616,000	$444,000	$236,500	$118,000	0
ONE-TIME STUDY						
Average Monthly Travel Time $\bar{T}' \mid \ell$ (minutes)	1076.2	1035.5	994.8	954.2	933.9	913.6
Percent Increase in Travel Associated Cost	18%	13%	9%	4%	2%	0%
Average Total Cost Per Interveiwer Per Month	$274.86	$268.13	$261.4	$254.68	251.32	$247.95
Percent Increase in Total Cost	11%	8%	5%	3%	1.5%	0%
Total Cost of the Study (over 27 months)	$955,000	$716,000	$477,000	$239,000	$120,000	0

1/ Note that these figures do not include any processing or analysis costs.

Appendix I. Derivation of Segment-to-Segment Travel Components for Interpenetrated Areas

A. Travel Time Component

For noninterpenetrated PSUs the segment-to-segment travel time component is taken to be:

$$T_{ssn} = (\lambda_1 s_1 - \lambda_2) d_1 r_1$$

To adjust this term to reflect the sample being interpenetrated, assume that x of the interviewer's s_1 segments are in another interviewer's areas and define:

$I = x/s_1$; proportion of an interviewer's assignment in another interviewer's area

and $N = 1 - x/s_1$; proportion of the interviewer's original assignment retained.

For interpenetrated areas this travel component can be obtained by:

$T_{ssi} = T_{ss}$ (due to segments in another area) + T_{ss} (due to segments in own area)

Consequently, T_{ssi} can be written as:

$$T_{ssi} = (\lambda_1(Is_1) - (I\lambda_2)) (d_1/\sqrt{I})\, r_1 + (\lambda_1(Ns_1) - (N\lambda_2)) (d_1/\sqrt{N}) r_1$$

where Is_1 and Ns_1 are the expected number of segments for each of the two areas, $I\lambda_2$ and $N\lambda_2$ are the average number of trips from home to each of the two areas,

and $d_1/\sqrt{I}$ and $d_1/\sqrt{N}$ are the average distances between segments in each of the two areas. A derivation of how these distances were obtained is given in Section B of this appendix.

By letting $N = 1-I$ the above simplifies to:

$$T_{ssi} = (\lambda_1 s_1 - \lambda_2)\, d_1\, r_1 \left(\sqrt{I} + \sqrt{1 - I}\right)$$

Table 7. Travel Time Components by Travel Strata by Type of Interpenetrated Area (in minutes)

Type of Interpenetrable Areas	Non-Interpenetrable PSUs					Multiple-Interviewer PSUs					Single-Interviewer PSUs				
Travel Time Component / Travel Strata	A	B	C	D	E	A	B	C	D	E	A	B	C	D	E
T_{ss}	455.8	473.3	346.3	383.6	306.3	$455.8f_1$	$473.3f_1$	$346.3f_1$	$383.6f_1$	$306.3f_1$	$455.8f_1$	$473.3f_1$	$346.3f_1$	$383.6f_1$	$306.3f_1$
T_{hs}	452.9	334.0	249.6	262.9	502.9	$452.9f_2$	$334.0f_2$	$249.6f_2$	$262.9f_2$	$502.9f_2$	$452.9f_{3A}$	$344.0f_{3B}$	$249.6f_{3C}$	$262.9f_{3D}$	$502.9f_{3E}$
T_{ws}	38.6	59.8	128.3	191.5	277.2	38.6	59.8	128.3	191.5	277.2	38.6	59.8	128.3	191.5	277.2
T	947.3	867.1	724.2	838.0	1086.4	--	--	--	--	--	--	--	--	--	--
Travel Strata Workloads	62	658	769	2136	5215	28,421	8,167	4,023	2,580	2,809	551	2,120	3,225	3,406	1,595
Type of Area Workloads	$W_n = 8,840$					$W_m = 45,990$					$W_s = 10,897$				
	$T_{ssn} = 341.9$ $T_{hsn} = 410.0$ $T_{wsn} = \underline{275.7}$ $T_h = 977.6$					$T_{ssi(m)} = 436.1\ (\sqrt{T} + \sqrt{T-T})$ $T_{hsi(m)} = 406.4\ (1 + 1)$ $T_{wsi(m)} = \underline{73.4}$ $T_{i(m)} =$					$T_{ssi(s)} = 302.3\ (\sqrt{T} + \sqrt{T-T})$ $T_{hsi(s)} = 317.5 + 1031.21$ $T_{wsi(s)} = \underline{152.0}$ $T_{i(s)} =$				

$$T = [(18,840)\ (977.6) + (45,990)\ T_{i(m)} + (10,897)\ T_{i(s)}]/65,727 \quad T_i = [(45,990)\ T_{i(m)} + (10,897)\ T_{i(s)}]/56,887$$

Definition of factors: $f_1 = \sqrt{T} + \sqrt{T-T}$ $\quad f_{3A} = 1 + 11.7\ 1$ $\quad f_{3C} = 1 + 3.4\ 1$ $\quad f_{3E} = 1 + 2.1\ 1$

$f_2 = (1 + 1)$ $\quad f_{3B} = 1 + 3.1\ 1$ $\quad f_{3D} = 1 + 1.9\ 1$

Note that the computations for the f_3 factors are shown in Table 2, in Appendix II.

Note that $W_i = W_m + W_s = 56,887$

Appendix I

To simplify notation in Table 7, f_1 is defined to be ($\sqrt{I} + \sqrt{1-I}$) and therefore,

$$T_{ssi} = (\lambda_1 s_1 - \lambda_2) d_1 r_1 f_1$$

The above assumptions appear to be quite adequate for multi-interviewer PSUs and single-interviewer PSUs and therefore T_{ssi} is applicable for both types of areas. This implies,

$$T_{ssi(m)} = T_{ssi(s)} = T_{ssi}$$

These overall segment-to-segment travel time components for multi- and single-interviewer PSUs were estimated by forming a weighted average of the five travel strata using travel stratum workloads as weights.

This resulted in the following estimates which are also given in Table 1.

$$\bar{T}_{ssi(m)} = 436.1 (\sqrt{I} + \sqrt{1-I})$$

and $$T_{ssi(s)} = 382.3 (\sqrt{I} + \sqrt{1-I})$$

B. Derivation of Distances

On page 112 in [5] Pielou gives the probability density function of the distribution of distances from a point to its nearest neighbor to be,

$$f(r) = 2\lambda r e^{\lambda r^2}$$

where r is the distance and λ is the mean number of points per circle of unit radius (i.e., a circle of area Π). This probability distribution assumes that the pattern of the points is random. The expected value of the distance r is also shown to be

$$E (r) = 1/2(\Pi/\lambda)^{\frac{1}{2}}$$

Appendix I

Now if points are taken to be segments and if it is assumed that interviewers travel to the nearest segment then the E(r) can be taken to be d_1, thus

$$d_1 = \frac{1}{2} (\pi/\lambda)^{\frac{1}{2}} \qquad [1]$$

and $\lambda = \pi/(2d_1)^2$; the mean number of segments per circle of area π .

Since an interviewer has s_1 segments, this implies that there are s_1/λ circles of area π in an interviewer's area. Therefore this area can be written as:

$$A = (s_1/\lambda)\,\pi$$

To find the segment-to-distance for $x = Is_1$ segments, assume that the total area remains fixed, which implies that the mean number of segments per circle of area π will change.

Let λ' be the new mean given x segments. Thus, the area given above can be written in terms of these parameters, which results in:

$$A = (x/\lambda')\pi = (s_1/\lambda)\,\pi$$

λ' can then be written in terms of λ as:

$$\lambda' = (x/s_1)\lambda$$

Using [1] above, the distance d_1', for an area containing x segments is:

$$d_1' = \frac{1}{2} (\pi/\lambda')^{\frac{1}{2}}$$

Appendix I.

Using $\lambda' = (x/s_1)\ \lambda$ gives

$$d_1' = \frac{1}{2}\ (\pi\ s_1/x\ \lambda)^{\frac{1}{2}} = \frac{r}{2}\ (\frac{\pi}{\lambda})^{\frac{1}{2}}\ (\frac{s_1}{x})^{\frac{1}{2}}$$

and since $d_1 = \frac{1}{2}\ (\frac{\pi}{\lambda})^{\frac{1}{2}}$ and $x/s_1 = I$

This results in:

$$d_1' = d_1/\sqrt{I}$$

Similarly an estimate of the distance between segments can be obtained for the $s_1 - x$ segments in the interviewers own area; which would give:

$$d_1' = d_1/\sqrt{I}\ ;\ \text{where } N = 1 - I$$

Appendix II Derivation of Home-To-Segment Travel Components for Interpenetrated Areas

As stated in the text this travel time component for noninterpenetrated PSUs is given by:

$$T_{hsn} = 2 \lambda_2 d_2 r_2$$

For interpenetrated areas this component is actually a function of the following two terms:

$T_{hsi} = T_{hs}$ (due to segments in own area) + T_{hs} (due to segments in another area)

Using N, I, $N\lambda_2$ and $I\lambda_2$ as defined in Appendix I this component can be written as:

$$T_{hsi} = 2 (N\lambda_2) d_2 r_2 + 2 (I\lambda_2) d_4 r_2$$

Where d_4 is the distance from the interviewer's home to segments in another interviewer's area.

For multiple interviewer PSUs a reasonable assumption is that d_4 can be taken to be $2d_2$, since we are assuming that only contiguous interviewer areas within a PSU will be paired and also that segments are uniform within a PSU. Therefore, this home-to segment travel time component for

Appendix II

multiple interviewer areas, $T_{hsi(m)}$, can be written as:

$$T_{hsi(m)} = 2\ (N\lambda_2)\ d_2\ r_2 + 2\ (I\lambda_2)\ (2d_2)\ r_2.$$

By letting N = 1 - I this simplifies to:

$$T_{hsi(m)} = 2\ d_2\ r_2\ \ \lambda_2\ (1 + I).$$

To simplify the notation in Table 7 f_2 is defined to be (1 + I) and

$$T_{hsi(m)} = 2\ d_2\ r_2\ \lambda_2\ f_2.$$

This gives the home-to-segment travel time components for the five multi-interviewer PSU strata to be:

$452.9f_2$, $334.0f_2$, $249.6f_2$, $262.9f_2$ and $502.9f_2$

for stratum A, B, C, D and E respectively.

Appendix II

The weighted average of these components, using the stratum workloads as weights, was then used to obtain the overall estimate for multiple interviewer PSUs which is,

$$T_{hsi(m)} = 406.4\ (I + I)$$ and is given in Table 7.

For single-interviewer PSUs, $2d_2$ may be a gross underestimate of d_4 since the "other area" would be in another PSU. A more appropriate value to use would be the distance between the centers of the paired PSUs. This creates a problem though when a single-interviewer from the multiple interviewer PSU was assumed to have traveled $2d_2$ miles from home to segments in the other area and will actually have traveled d_4 miles. To easily adjust for this, d_4 (for single interviewer PSUs paired with multi-interviewer PSUs) was adjusted by a factor of $d_4/2d_2$ to account for the increased travel not included in the term $T_{hsi(m)}$. Thus for single interviewer PSUs we have:

$$T_{hsi(s)} = 2\ (N\lambda_2)\ d_2\ r_2 + 2\ (I\lambda_2)\ d_4\ r_2$$

where d_4 is the distance between centers of paired PSUs and includes the $d_4/2d_2$ adjustment factor.

Appendix II.

Again letting N = 1 - I the above simplifies to:

$$T_{hsi(s)} = 2 \lambda_2 d_2 r_2 [1 + (d_4/d_2 - 1) I] .$$

Since d_4/d_2 will depend on the particular travel strata (i.e, A,B,C,D or E), a factor was determined for each of these strata and is defined to be:

$$f_{3i} = 1 + (\frac{d_{4i}}{d_{2i}} - 1) I \text{ for } i = A, B, C, D \text{ or } E$$

The computations for these factors are given in Table 8. Therefore, the home-to-segment travel times for the five single-interviewer PSU strata are given by:

$$452.9f_{3A},\ 334.0f_{3B},\ 249.6f_{3C},\ 262.9f_{3D} \text{ and } 502.9f_{3E}$$

for strata A, B, C, D, and E respectively.

The weighted average of these components, using the stratum workloads as weights, was then used to obtain the overall estimate for single-interviewer PSUs which is

$$T_{hsi(s)} = 317.5 + 1031.2 I.$$

This term is also given in Table 7.

Table 8. Computations for the f_3 factors

Single Interviewer Travel Stratum	d_{2i} 1/	Σd_4 2/	wkld	$d_{4i} = \frac{\Sigma d_4}{wkld}$	$\frac{d_{4i}}{d_{2i}} - 1$	$f_3 = (1 + (\frac{d_{4i}}{d_{2i}} - 1) I)$
A	6.1	42,677	551	77.5	11.7	1 + 11.7I = f_{3A}
B	12.1	106,280	2,120	50.1	3.1	1 + 3.1 I = f_{3B}
C	12.0	171,625	3,225	53.2	3.4	1 + 3.4 I = f_{3C}
D	15.8	156,171	3,406	45.9	1.9	1 + 1.9 I = f_{3D}
E	16.9	83,408	1,595	52.3	2.1	1 + 2.1 I = f_{3E}

1/ Values contained in the memorandum, "Relative Size of the Components of CPS Travel Time," dated March 12, 1981 from Blass to Kniceley.

2/ These were obtained by direct measurement and includes the additional $\frac{d_4}{2d_2}$ factor when necessary.

Appendix III. Computations of Travel Times

Case 1: $\bar{T}$; expected travel time in minutes per interviewer when no sample is interpenetrated. Assume I = 0.

Non-Interpenetrable PSUs	Multi-Interviewer PSUs	Single Interviewer PSUs
$\bar{T}_{ssn}$ = 341.9	$\bar{T}_{ssi(m)}$ = 436.1	$\bar{T}_{ssi(s)}$ = 382.3
$\bar{T}_{hsn}$ = 410.0	$\bar{T}_{hsi(m)}$ = 406.4	$\bar{T}_{hsi(s)}$ = 317.5
$\bar{T}_{wsn}$ = 225.7	$\bar{T}_{wsi(m)}$ = 73.4	$\bar{T}_{wsi(s)}$ = 152.0
$\bar{T}_{n}$ = 977.6	$\bar{T}_{i(m)}$ = 915.9	$\bar{T}_{i(s)}$ = 851.8

$$\bar{T} = [\ (8,840)(977.6) + (45,990)(915.9) + (10,897)(851.8)\ ] / 65,727$$

$$\bar{T} = 913.6 \text{ minutes}$$

Appendix III.

Case 2: $\bar{T}_{i(s)}$; expected travel time in minutes for an interviewer in a single-interviewer PSU, included only one time in the study. Assume I = .2222, the average value of I over the 27 months.

$$\bar{T}_{ssi(s)} = 382.3\ (\sqrt{.2222} + \sqrt{.7778})$$

$$\bar{T}_{hsi(s)} = 317.5 + 1031.2(.2222)$$

$$\bar{T}_{wsi(s)} = \underline{152.0}$$

$$\bar{T}_{i(s)} = 1216.0 \text{ minutes}$$

Case 3: $\bar{T}_{i(m)}$; expected travel time in minutes for an interviewer in a multi-interviewer PSU, included only one time in the study Again assume I = .2222.

$$\bar{T}_{ssi(m)} = 436.1\ (\sqrt{.2222} + \sqrt{.7778})$$

$$\bar{T}_{hsi(m)} = 406.4\ (1.2222)$$

$$\bar{T}_{wsi(m)} = \underline{73.4}$$

$$\bar{T}_{i(m)} = 1160.3 \text{ minutes}$$

Appendix III.

Case 4: $\bar{T}_i$; expected travel time in minutes for an interviewer in any interpenetrable area included only one time in the study. Again assume I = .2222.

$$\bar{T}_i = [\ (45{,}990)(1160.3) + (10{,}897)(1216.0)\ ] / 56{,}887$$

$$\bar{T}_i = 1170.7 \text{ minutes}$$

Case 5: $\bar{T}'|\ell$; average monthly travel time for all interviewers for the 27 affected months when ℓ pairs are selected for the study. The general form for estimating this travel time is given by:

$$\bar{T}'|\ell = \frac{W_n \bar{T}_n + W_i\{(\ell /L)(\bar{T}_i) + (1-\ell/L)[(W_m\bar{T}_{i(m)} + W_s \bar{T}_{i(s)}/W_i]\}}{W_n + W_i}$$

where W_n and W_i are workloads (i.e., housing units) for non-interpenetrable and interpenetrable areas, respectively.

W_m and W_s are workloads for multi- and single-interviewer PSUs respectivel in interpenetrable areas (i.e., $W_m + W_s = W_i$).

$\bar{T}_n$ and $\bar{T}_i$ are the expected monthly travel times for non-interpenetrated and interpenetrated interviewers respectively for the 27 affected months.

$\bar{T}_{i(m)}$ and $\bar{T}_{i(s)}$ are the expected monthly travel times for interviewers in multi- and single- interviewer interpenetrable areas respectively; wher <u>no</u> sample is interpenetrated.

L is the expected number of interviewer pairs that could be interpenetrate

ℓ in the number of interviewer pairs selected for the study.

Appendix III.

Case 5: (continued)

From the data used in this report $T'|\ell$ can be expressed as:

$$T'|\ell = \{(8{,}840)(977.6) + (56{,}887)\{(\ell/568.87)(1170.7)$$
$$+ (1-\ell/568.87)[(45{,}990)(915.9) + 10{,}897)(851.8))/56{,}887]\}\}/(8{,}840 + 56{,}887)$$

where the W's can be obtained from Table 7, the values 977.6, 915.9 and 851.8 were obtained in Case 1 of this appendix and the value 1170.7 was also obtained in this appendix (i.e., case 4).

Appendix IV. Determination of Travel Time Components Given ℓ Pairs Selected for Year i

For a given year the average proportion of an interviewer's assignment in another interviewer's area τ, will depend on the year(s) in which the interviewer was selected for Pattern I. For example in year i = 1985, the value of τ will depend on whether the interviewer was selected for the study in 1985, 1984 and 1983, since an interviewer selected for the study in one year will still have sample interpenetrated 27 months later. Thus to determine the value of τ for year i, the joint probability density function of the random vector (X_i, X_{i-1}, X_{i-2}) was obtained and is denoted by $p(X_i, X_{i-1}, X_{i-2})$ where

$$X_j = \begin{cases} 0 \text{ if the pair was not selected in year } j \\ 1 \text{ if the pair was selected in year } j \end{cases}$$

Define $X. = \sum_{j=i-2}^{i} x_j$; Then for $X. \neq 0$, X. represents the number of years in which the pair has been included and consequently has interpenetrated sample in year i.

The joint density function of (X_i, X_{i-1}, X_{i-2}) is

$$p(X_i, X_{i-1}, X_{i-2}) = \left(\frac{\ell}{L}\right)^{X.} \left(1-\frac{\ell}{L}\right)^{3-X.}$$

for X. = 0, 1, 2, or 3

where L is the total possible interpenetrable pairs. This assumes that the selection of ℓ pairs is done independently each year.

For given year the average proportion of an interviewer's assignment in another interviewer's area, $\tau(X_i, X_{i-1}, X_{i-2})$, will depend on the random vector (X_i, X_{i-1}, X_{i-2}).

This value can be obtained by:

$$\tau(X_i, X_{i-1}, X_{i-2}) = \sum_{j=i-2}^{i} \tau_{Xj}$$

Appendix IV.

$$\text{where } T_{X_j} = \begin{cases} X_i\ (.2188) \text{ for } j = i \\ X_{i-1}\ (.2500) \text{ for } j = i - 1 \\ X_{i-2}\ (.0312) \text{ for } j = i - 2 \end{cases}$$

Note that .2188 is the average value of I for the 1st 12 months of interpenetration, .2500 for the average value of I for the 2nd 12 months of interpenetration and .0312 is the average value of I for the 3rd 12 months of interpenetration. Note that actually for 9 of the 3rd 12 months no sample is interpenetrated.

These average values of I can be obtained from the following table:

Table 9. I by Month in Study

Month	I (1st Year)	I (2nd Year)	I (3rd Year)
1	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{3}{16}$
2	$\frac{2}{16}$	$\frac{4}{16}$	$\frac{2}{16}$
3	$\frac{3}{16}$	"	$\frac{1}{16}$
4	$\frac{4}{16}$	"	0
5	"	"	"
6	"	"	"
7	"	"	"
8	"	"	"
9	"	"	"
10	"	"	"
11	"	"	"
12	"	"	"
Average I	.2188	.2500	.0312

Appendix IV.

The following table gives a summary of the distribution of all possible events and their associated probability and the resulting value of I.

Table 10. Summary of Events

(x_i, x_{i-1}, x_{i-2})	$T(x_i, x_{i-1}, x_{i-2})$	$x.$	$p(x_i, x_{i-1}, x_{i-2})$
(1, 1, 1)	.5000	3	$\left(\frac{\ell}{L}\right)^3$
(1, 1, 0)	.4688	2	$\left(\frac{\ell}{L}\right)^2 (1 - \frac{\ell}{L})$
(1, 0, 1)	.2500	2	$\left(\frac{\ell}{L}\right)^2 (1 - \frac{\ell}{L})$
(0, 1, 1)	.2812	2	$\left(\frac{\ell}{L}\right)^2 (1 - \frac{\ell}{L})$
(0, 0, 1)	.0312	1	$\left(\frac{\ell}{L}\right) (1 - \frac{\ell}{L})^2$
(0, 1, 0)	.2500	1	$\left(\frac{\ell}{L}\right) (1 - \frac{\ell}{L})^2$
(1, 0, 0)	.2188	1	$\left(\frac{\ell}{L}\right) (1 - \frac{\ell}{L})^2$
(0, 0, 0)	0	0	$(1 - \frac{\ell}{L})^3$

The travel time components dependent on I are:

$$T_{ssi(m)},\ T_{ssi(s)},\ T_{hsi(m)} \text{ and } T_{hsi(s)}$$

These components can be estimated given that ℓ pairs are selected by:

$$T_{ssi(m)}|\ell = \Sigma\, p(x_i, x_{i-1}, x_{i-2})\ (T_{ssi(m)} \mid T{=}I\ (x_i, x_{i-1}, x_{i-2}))$$

$$T_{hsi(m)}|\ell = \Sigma\, p(x_i, x_{i-1}, x_{i-2})\ (T_{hsi(m)} \mid T{=}I(x_i, x_{i-1}, x_{i-2}))$$

$$T_{hsi(s)}|\ell = \Sigma\, p(x_i, x_{i-1}, x_{i-2})\ (T_{hsi(s)} \mid T{=}I\ (x_i, x_{i-1}, x_{i-2}))$$

$$T_{ssi(s)}|\ell = \Sigma\, p(x_i, x_{i-1}, x_{i-2})\ (T_{ssi(s)} \mid T{=}I\ (x_i, x_{i-1}, x_{i-2}))$$

Note that the above Σ's are over all possible (x_i, x_{i-1}, x_{i-2}) events.

Appendix IV.

To obtain the overall expected travel time per interviewer given that ℓ pairs are selected $(\bar{T}|\ell)$, simply replace the terms $\bar{T}_{ssi(m)}$, $\bar{T}_{ssi(s)}$, $\bar{T}_{hsi(m)}$ and $\bar{T}_{hsi(s)}$ in Table 7 with appropriate formula given above and proceed as before. An example of the calculations involved in obtaining $\bar{T}|\ell = 400$ is given in Table 11. Note that a more precise estimate could be obtained by averaging the above components for each month. Use of $\bar{T}$, however, considerably simplifies the computation and provides an extremely reasonable approximation. For example, in Table 11, $\bar{T}|\ell=400$ is estimated as 1210.4. Use of a more precise estimator would give 1207.84.

Table 11. Example of Calculations for T|ℓ = 400

$,X_{i-2})$	$I(X_i,X_{i-1},X_{i-2})$ (2)	$P(X_i,X_{i-1},X_{i-2})$ (3)	$P(X_i,X_{i-1},X_{i-2})$ for ℓ =400 (4)	$\bar{T}_{ssi(m)}\|I=I(X_i,X_{i-1},X_{i-2})$ (5)	(4)X(5) (6)	$\bar{T}_{hsi(m)}\|I=I(X_i,X_{i-1},X_{i-2})$ (7)	(7)X(4) (8)	$\bar{T}_{ssi(s)}\|I=I(X_i,X_{i-1},X_{i-2})$ (9)	(9)X(4) (10)	$\bar{T}_{hsi(s)}\|I=I(X_i,X_{i-1},X_{i-2})$ (11)	(11)X((12
1)	.5000	$(\ell/L)^3$	.348	616.7	214.6	609.6	212.1	540.7	188.1	833.1	289.9
0)	.4688	$(\ell/L)^2(1-\ell/L)$	.147	616.4	90.6	596.9	87.8	540.4	79.4	800.9	117.7
, 1)	.2500	$(\ell/L)^2(1-\ell/L)$	.147	595.7	87.6	508.0	74.7	522.2	76.8	575.3	84.6
, 1)	.2812	$(\ell/L)^2(1-\ell/L)$	.147	601.0	88.3	520.7	76.5	526.8	77.4	607.5	89.3
, 1)	.0312	$(\ell/L)(1-\ell/L)^2$	.062	506.1	31.4	419.1	26.0	443.8	27.5	349.7	21.7
, 0)	.2500	$(\ell/L)(1-\ell/L)^2$	.062	595.7	36.9	508.0	31.5	522.2	32.4	575.3	35.7
, 1)	.2188	$(\ell/L)(1-\ell/L)^2$	.062	589.4	36.5	495.3	30.7	516.7	32.0	543.1	33.7
, 0)	0	$(1-\ell/L)^3$	.026	436.1	11.3	406.4	10.6	382.3	9.9	317.5	8.3

al L = 568.87

li-Interviewer Components

Column (6) gives $\bar{T}_{ssi(m)}|\ell=400$ = 597.2

Column (8) gives $\bar{T}_{hsi(m)}|\ell=400$ = 549.9

ves $\bar{T}_{i(m)}|\ell=400$ = 597.2 + 549.9 + 73.4 = 1220.5

Single-Interviewer Components

Sum of Column (10) gives $\bar{T}_{ssi(s)}|\ell=400$ = 523.5

Sum of Column (12) gives $\bar{T}_{hsi(s)}|\ell=400$ = 680.9

This gives $\bar{T}_{i(s)}|\ell=400$ = 523.5 + 680.9 + 152.0 = 1356.4

0 For all Areas:

ℓ =400 = [(8,840)(977.6) + (45,990)(1220.5) + (10,897)(1356.4)]/65,727

=400 = 1210.4 which results in a percent increase of $(1210.4/913.6-1)(100)$ = 32%

Appendix V. Yearly Estimates of Interviewer Variance [1]

1. INTRODUCTION

Suppose an annual estimate of interviewer variance is made as the average of the estimate for each month; i.e.,

$$\sigma^2_{b,yr}(1) = \frac{1}{12}\sum_{\alpha=1}^{12}\sigma^2_{b\alpha}(1),$$

where $\sigma^2_{b\alpha}(1)$ is the usual estimate of interviewer variance for the α^{th} month as discussed in [2]. For simplicity, $\sigma^2_{b\alpha}(1)$ and $\sigma^2_{b,yr}(1)$ will be denoted from here on by $\sigma^2_{b\alpha}$ and $\sigma^2_{b,yr}$. Let $y_\alpha = ((y_{\alpha\gamma tjs}))$ denote the s^{th} observation recorded by the j^{th} interviewer in the t^{th} enumeration area (EA) of the γ^{th} block in the α^{th} month. Then

$$\hat{\sigma}^2_{b\alpha} = \sum_{i=1}^{\mathcal{l}+1} \lambda_{oi}\, \underline{y}_\alpha' \, A_i \, \underline{y}_\alpha,$$

where $\underline{\lambda}_o$, A_i are defined in (4.3) and (4.7) of [2] and $\mathcal{l}$ is the number of EA's. For the balanced interpenetration scheme where $\mathcal{l} = k\ell$, ℓ is the number of interpenetrated blocks, k is the number of EA's per block, and m is the size of each EA, we have from Appendix B of [2].

$$A_1 = \underline{1}_{ff} \otimes (I_k - \frac{1}{k}\underline{1}_{kk}) \otimes I_{\mathcal{l}} \qquad (1)$$

$$A_i = (\underline{0}_m, \ldots, I_m - \frac{1}{m}\underline{1}_{mm}, \ldots, 0_m), \; i=2,\ldots,\mathcal{l}+1, \qquad (2)$$

with the non-zero matrix in the $(i-1)^{th}$ position, and

$$\lambda_o = \frac{1}{(km-2k+1)\ell m}\left[\frac{k(m-1)}{(k-1)m}, -1, \ldots, -1\right]. \qquad (3)$$

In Section 1, we derive the variance of $\hat{\sigma}^2_{b,yr}$. This is used in Section 2

[1] Reproduced from the memo from Lynn Stokes to Paul Biemer, "CPS Interview Variance Estimates, August 9, 1983.

to design an interpenetration scheme for CPS which achieves the desired precision for the annual estimate of σ_b^2.

1. Var $\hat{\sigma}^2_{b,yr}$

The monthly estimates $\sigma^2_{b\alpha}$ are correlated since we assume:

(a) the interviewers remain the same from month to month

and

(b) each respondent remains in sample for c months before rotating out.

Thus

$$\text{Var } \hat{\sigma}^2_{b,yr} = \frac{1}{144}\left[\sum_{\alpha=1}^{12} \text{Var } \hat{\sigma}^2_{b\alpha} + 2 \sum_{\alpha<\beta}\sum \text{Cov}(\hat{\sigma}^2_{b\alpha}, \hat{\sigma}^2_{b\beta})\right]. \quad (4)$$

From (5.3) of [2] we see that

$$\text{Var } \hat{\sigma}^2_{b\alpha} = \frac{2}{\ell(k-1)}\left[\sigma_b^4 + \frac{2}{m}\sigma_b^2 \sigma_e^2 + \frac{k(m-1)}{m^2(km-2k+1)}\sigma_e^4\right]. \quad (5)$$

Note that

$$\text{Cov}(\hat{\sigma}^2_{b\alpha}, \hat{\sigma}^2_{b\beta}) = \text{Cov}\left(\sum_{i=1}^{\ell+1} \lambda_{oi}\, \underset{\sim}{y}_\alpha' A_i \underset{\sim}{y}_\alpha, \sum_{j=1}^{\ell+1} \lambda_{oj}\, \underset{\sim}{y}_\beta' A_j \underset{\sim}{y}_\beta\right)$$

$$= \sum_i \sum_j \lambda_{oi} \lambda_{oj} \text{Cov}(\underset{\sim}{y}_\alpha' A_i \underset{\sim}{y}_\alpha, \underset{\sim}{y}_\beta' A_j \underset{\sim}{y}_\beta) \quad (6)$$

and

$$\text{Cov}(\underset{\sim}{y}_\alpha' A_i \underset{\sim}{y}_\alpha, \underset{\sim}{y}_\beta' A_j \underset{\sim}{y}_\beta) = \text{Cov}(\underset{\sim}{y}_{\alpha\beta}' A_{i1} \underset{\sim}{y}_{\alpha\beta}, \underset{\sim}{y}_{\alpha\beta}' A_{j2} \underset{\sim}{y}_{\alpha\beta})$$

$$= 2\,\text{tr}(A_{i1} \Sigma_{\alpha\beta} A_{j2} \Sigma_{\alpha\beta})$$

$$= 2\,\text{tr}(A_i \Sigma_{\alpha\beta 12} A_j \Sigma_{\alpha\beta 12}). \quad (7)$$

where $A_{i1} = \begin{pmatrix} A_i & 0 \\ 0 & 0 \end{pmatrix}$, $A_{i2} = \begin{pmatrix} 0 & 0 \\ 0 & A_i \end{pmatrix}$, $\underset{\sim}{y}_{\alpha\beta}' = (\underset{\sim}{y}_\alpha', \underset{\sim}{y}_\beta')$, and $\Sigma_{\alpha\beta} = \text{Cov}(\underset{\sim}{y}_{\alpha\beta})$

$$= \begin{bmatrix} \Sigma_{\alpha\beta 11} & \Sigma_{\alpha\beta 12} \\ \Sigma_{\alpha\beta 12} & \Sigma_{\alpha\beta 22} \end{bmatrix};$$

We assume that

(a) two observations made by the same interviewer on different respondents in different months have covariance σ_b^2; i.e.,

$Cov(y_{\alpha\gamma tjs}, y_{\beta\gamma tjs'}) = \sigma_b^2$.

and

(b) two observations made by the same interviewer on the same respondents in different months have covariance $\rho_c\sigma_e^2 + \sigma_b^2$; i.e.,

$Cov(y_{\alpha\gamma tjs}, y_{\beta\gamma tjs}) = \rho_c\sigma_e^2 + \sigma_b^2$ if the respondent is the same individual in months α and β.

Thus,

$$\Sigma_{\alpha\beta 12} = \mathrm{Diag}(D,\ldots,D)$$

where
$$D_{mk \times mk} = \begin{bmatrix} \sigma_b^2 B + \rho_c\sigma_e^2 R & \sigma_b^2 B & \cdots & \sigma_b^2 B \\ \vdots & \vdots & & \vdots \\ \sigma_b^2 B & \sigma_b^2 B & \cdot & \sigma_b^2 B + \rho_c\sigma_e^2 R \end{bmatrix},$$

$$B_{m \times m} = \begin{bmatrix} \underline{1}_{ff} & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \underline{1}_{ff} \end{bmatrix}, \quad R_{m \times m} = \begin{bmatrix} R_o & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & R_o \end{bmatrix}, \quad R_o = \begin{bmatrix} 0 & 0 \\ 0 & I_{\delta f} \end{bmatrix},$$

where $f = m/k$ and δ = the proportion of respondents which are common to the sample in the α^{th} and β^{th} months.

One can show that

(a) $tr(A_1\,\Sigma_{\alpha\beta 12})^2 = \ell m^2(k-1)(m\sigma_b^2 + \delta\rho_c\sigma_e^2)^2$

(b) $tr(A_1\,\Sigma_{\alpha\beta 12}\,A_j\,\Sigma_{\alpha\beta 12})$

$$= \frac{m}{k}(k-1)[(m\sigma_b^2 + \delta\rho_c\sigma_e^2)^2 + \delta(1-\delta)\,\rho_c^2\sigma_e^4],\ j=2,\ldots,\ell+1$$

(c) $tr(A_i \Sigma_{\alpha\beta 12})^2 = \frac{m^2}{k^2}(k-1)\ \sigma_b^4 + 2\delta\ \frac{m}{k}(k-1)\ \sigma_b^2 \rho_c \sigma_e^2$

$+ \sigma_e^4\ [m\delta - \delta(2-\delta)], \qquad i=2,\ldots,\ell+1$

(d) $tr(A_i \Sigma_{\alpha\beta 12} A_j \Sigma_{\alpha\beta 12}) = \begin{cases} 0 & \text{for } i,j = 2,\ldots,\ell+1; i\neq j; \text{ if } (i-1)^{th} \text{ and } (j-1)^{th} \text{ EA's are in different blocks} \\ \frac{m^2}{k^2}(k-1)\ \sigma_b^4 & \text{for } i,j = 2,\ldots,\ell+1; i\neq j; \text{ if } (i-1)^{th} \text{ and } (j-1)^{th} \text{ EA's are in the same block.} \end{cases}$

Thus from (1) through (3), (6), and (7), we have

$$Cov(\sigma_{b\alpha}^2, \sigma_{b\beta}^2) = \frac{2}{\ell}\ \{ \frac{\sigma_b^4}{(k-1)} + \frac{2\delta}{m(k-1)}\ \sigma_b^2 \rho_c \sigma_e^2$$

$$+ \frac{\delta k \rho_c^2 \sigma_e^4}{[k(m-2)+1]^2 m^2}\ [\frac{k}{k-1}\ (m-1)^2\ \delta - (m-\delta)]\}. \qquad (8)$$

When k=2,

$$Cov(\sigma_{b\alpha}^2, \sigma_{b\beta}^2) = \frac{2}{\ell}\ \{\sigma_b^4 + \frac{2\delta}{m}\ \sigma_b^2 \rho_c \sigma_e^2$$

$$+ \frac{2\delta}{m^2(2m-3)^2}\ [2(m-1)^2\ \delta - (m-\delta)]\rho_c^2 \sigma_e^4\}.$$

When c = 4, as in the CPS, we have:

(a) $\delta = 3/4$ if $|\alpha - \beta| = 1$

(b) $\delta = 1/2$ if $|\alpha - \beta| = 2$

(c) $\delta = 1/4$ if $|\alpha - \beta| = 3$

(d) $\delta = 0$ if $|\alpha - \beta| \geq 4$

In that case, from (4),

$$\text{Var } \hat{\sigma}^2_{b,yr} = \frac{1}{144} [12 \text{ Var } \hat{\sigma}^2_{b\alpha} + 2(11) \text{ Cov}(\hat{\sigma}^2_{b\alpha}, \hat{\sigma}^2_{b,\alpha+1}) + 2(10) \text{ Cov}(\hat{\sigma}^2_{b\alpha}, \hat{\sigma}^2_{b,\alpha+2}) + 2(9) \text{ Cov}(\hat{\sigma}^2_{b\alpha}, \hat{\sigma}^2_{b,\alpha+3}) + 72 \text{ Cov}(\hat{\sigma}^2_{b\alpha}, \hat{\sigma}^2_{b,\alpha+4})]. \quad (9)$$

2. ANNUAL ESTIMATION OF σ^2_b in CPS

In this section, we determine the number of interpenetrated pairs of EA's needed to achieve a coefficient of variation of 0.2 for the CPS estimator $\sigma^2_{b,yr}$. The variance of $\sigma^2_{b,yr}$ given by (9) was based on the assumption of independence of the observations made in each month. So that this assumption will be more nearly met, let $\underline{y}_\alpha$ denote the vector of segment, rather than individual or household, observations. Thus if the segment variable is unemployment rate, then $\sigma^2_e = \frac{PQ}{\bar{s}}$ deff • BSEG where $\bar{s}$ is the average segment size, deff is the design effect, BSEG is the proportion of the within primary variance due to the between segment variability. P = overall proportion unemployed, and Q = 1-P.

The coefficient of variation, $CV = \left(\frac{\text{Var } \sigma^2_{b,yr}}{\sigma^4_b}\right)^{1/2}$, is a function of the ratio

$$\frac{\sigma^2_e}{\sigma^2_b} = \frac{PQ}{\sigma^2_b} \frac{(\text{deff})(\text{BSEG})}{\bar{s}} = (\rho_b^{-1} - 1) \frac{(\text{deff})(\text{BSEG})}{\bar{s}}, \text{ where } \rho_b = \frac{\sigma^2_b}{\sigma^2_b + PQ}.$$

Table 1 below shows the value of ℓ required to achieve CV = .2 and .1 for $.001 < \rho_b < .05$; $.5 < \rho_c < .7$, and where deff = 1.3, BSEG = .1, $\bar{s}$ = 4.8, and m = 12. These parameters were chosen to approximate those of the CPS, where interviewers' assignments average about 12 segments of about 4 households each. The average number in the labor force per household is assumed to be 1.2, yielding $\bar{s}$ = 4.8. Correlations between the unemployment rates based on the overlapping samples have been estimated from CPS to be between .5 and .7, depending on the number of months apart (personal communication from SMD). This correlation includes the between segment component, but we assume it to be small in comparison with the within-segment and within-household components.

Table 12.--Number of Interpenetrated Pairs Needed to Achieve CV = .2 (and .1)

BSEG = 10%, Deff = 1.3, $\bar{s}$ = 4.8, m = 12

ρ_b	ρ_c = .5	ρ_c = .6	ρ_c = .7
.001	124 (494)	132 (528)	141 (565)
.005	60 (239)	61 (244)	62 (248)
.01	55 (218)	55 (220)	56 (222)
.02	52 (209)	52 (210)	53 (211)
.05	51 (203)	51 (204)	51 (204)

REFERENCES

[1] Biemer, P. P. (1979), "An Improved Procedure for Estimating the Components of Response Variance in Complex Surveys," Proceedings of the ASA, Section on Survey Research Methods.

[2] Biemer, P.P. and Stokes, S.L. (1984), "Optimal Design of Interviewer Variance Experiements in Complex Surveys," Journal of the American Statistical Association, 80.

[3] Memorandum for Maurice Kniceley from Richard F. Blass (March 12, 1981). Subject: Relative Size of the Components of CPS Travel Time.

[4] Bureau of the Census. Evaluation and Research Program of the 1970 Census of Population and Housing: Enumerator Variance in the 1970 Census. PHC(E)-13.

[5] Pielou, E. C. (1969), "An Introduction to Mathematical Ecology", Wiley-Interscience, N.Y.