THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

A Composite Estimation for SIPP: A Preliminary Report

No. 18

Rameswar Chakrabarty

U.S. Census Bureau

June 1986

U.S. Department of Commerce U.S. CENSUS BUREAU

THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

A COMPOSITE ESTIMATION FOR SIPP
A PRELIMINARY REPORT

No. 8610 18

Rameswar P. Chakrabarty

SIPP Working Paper #18

June 1986

U.S. Department of Commerce BUREAU OF THE CENSUS

Acknowledgements

This paper was prepared by Rameswar P. Chakabarty, Statistical Research Division, Bureau of the Census.

The author would like to thank Lawrence Ernst and Nash Monsour, Jr., Statistical Research Division; Daniel Kasprzyk, Population Division; and Gary Shapiro and Raj Singh, Statistical Methods Division.

Suggested Citation

Chakrabarty, Rameswar P. "Composite Estimation for SIPP: A Preliminary Report," SIPP Working Paper Series No. 8610. Washington, D.C.: U.S. Bureau of the Census, 1986.

TABLE OF CONTENTS

1. Introduction

A. Survey Design 1
B. Estimation. 2

2. Annual Estimates. 6

3. Quarterly Estimates 9

4. Correlation and Robustness of Estimators. 13

5. Consistency Between Quarterly and Annual Estimates. 13

6. Other Approaches. 14

7. Time-in-Sample Bias 15

8. Within Interview Correlated Response Bias 16

9. Composite Estimation for Longitudinal Reports 16

10. Integration of Cross-Sectional and Longitudinal Estimation. . . . 17

11. Summary

A. General Estimates 24
B. Quarterly Estimates 24
C. Annual Estimate and Consistency Between Quarterly and Annual Estimates. 26
D. Further Research Needed 26

References 28

ATTACHMENTS

I. Composite Estimation in the CPS. 30

II. Data Structure for Annual Estimates. 39

III. Data Structure for Quarterly Estimates 41

Composite Estimation for SIPP
A Preliminary Report

1. Introduction

A. Survey Design

The Survey of Income and Program Participation (SIPP) will undoubtedly become a major source of data on the economic situation of households, families and individuals in the United States. The SIPP began in October 1983 as an ongoing survey program of the Bureau of the Census with a sample of approximately 25000 "designated" housing units selected to represent the noninstitutional population of the United States. Each sample household will be interviewed once every four months for 2 1/2 years to produce sufficient data for longitudinal analyses. The reference period is the four months preceeding the interview; for example, in October, the reference period is June through September. When the household is interviewed again in February, it is October through January. In February 1985 and every February thereafter, a new slightly smaller sample will be introduced. This overlapping panel design will enhance estimates of levels, and change, particularly year to year change. To facilitate field operations, each panel is divided into four approximately equal subsamples called rotation groups; one rotation group will be interviewed in a given month. Thus, one cycle or "wave" of interviewing generally takes four consecutive months. Figure 1 provides an illustration of the relationship between waves, rotation groups, interview months, and reference periods after the sample becomes fully operational in 1985. For further details of the design of the survey and plans for publication of reports see Nelson, McMillan and Kasprzyk (1984).

B. Estimation

The analysis of the SIPP data has two major aspects, cross-sectional and longitudinal analyses. This report primarily deals with estimatin for cross-sectional reports. Estimation of quarterly levels (monthly averages for a quarter) annual levels and year to year change in quarterly and annual levels are of major interest for cross-sectional reports. Composite estimation of monthly levels and changes from month to month or changes in monthly levels one year apart are of secondary importance and publication of such statistics is not contemplated now.

For the current occasion (a quarter or a year) the usual ratio estimator uses data only for the current occasion. This estimator is the result of four steps in the estimation procedure. These steps are (1) computation of simple unbiased estimates, (2) noninterview adjustment, (3) first stage ratio adjustment and (4) second stage ratio adjustment. In this report, the ratio estimator will be called 'simple' or 'elementary' (uncomposited) estimator.

A composite estimator makes use of data from the current as well as previous occasion(s). Different composite estimators employ different methods of combining data from the current and previous occasion(s).

Figure 1. SIPP Interview Months and Reference Periods

Wave	Rotation Group	Interview Month	Reference Periods																			
							1st Qtr Date			2nd Qtr Date			3rd Qtr Date			4th Qtr Date						
1	1	JAN	SEP	OCT	NOV	DEC																
	2	FEB		OCT	NOV	DEC	JAN															
	3	MAR			NOV	DEC	JAN	FEB														
	4	APR				DEC	JAN	FEB	MAR													
2	1	MAY					JAN	FEB	MAR	APR												
	2	JUN						FEB	MAR	APR	MAY											
	3	JUL							MAR	APR	MAY	JUN										
	4	AUG								APR	MAY	JUN	JUL									
3	1	SEP									MAY	JUN	JUL	AUG								
	2	OCT										JUN	JUL	AUG	SEP							
	3	NOV											JUL	AUG	SEP	OCT						
	4	DEC												AUG	SEP	OCT	NOV					
4	1	JAN													SEP	OCT	NOV	DEC				
	2	FEB														OCT	NOV	DEC	JAN			
	3	MAR															NOV	DEC	JAN	FEB		
	4	APR																DEC	JAN	FEB	MAR	
	1	MAY																	JAN	FEB	MAR	APR

It is well-known that whenever a survey has some overlap of samples over occasions a composite estimator usually reduces the variance of estimates of level and particularly of change. For example, the optimal composite estimator reduces the variance of the ratio estimator by about 32 percent when estimating changes in civilian labor force from the CPS (Breau and Ernat 1983). Composite estimators are specially effective in variance reduction when the correlation is high between responses from the same sample unit for consecutive occasions. The correlation between family incomes (adjusted gross incomes) reported to the IRS in 1979 and 1980 was found to be .80 to .87. The same results were also obtained between 1980 and 1981 family incomes. This indicates that the correlation between year to year incomes is high and also, stable (Ponikowski and Tadros, 1984). Thus, the potential for improving the reliability of SIPP estimates by appropriate composite estimation is considerable. It is appropriate to mention at this point that the variance reduction by composite estimation is attained at basically little or no additional cost because no extra field work is involved. This report outlines the results of research done so far in the composite estimation for SIPP cross-sectional reports.

From our analysis of the SIPP rotation pattern and data structure (see Attachments II and III), and research in composite estimation we have determined that the AK composite estimator (Huang and Ernst, 1981) currently used in CPS is applicable only to CPS-type rotation patterns (4-8-4 or 3-9-3). This estimator cannot be used for SIPP. The CPS 4-8-4 rotation pattern provides a simple and uniform data structure from month to month and year to year. The SIPP rotation pattern, on the other hand, creates a data structure that varies from month to month and quarter to quarter (see Attachment III). This lead to a comprehensive review of literature and research on composite estimation. We have determined types of composite estimators appropriate for SIPP annual and quarterly estimates (for persons, households or families) and these are discussed in the sequel.

The efficiency of a composite estimator for a characteristic depends on two key parameters--the proportion of the sample matched with previous occasion(s) and the correlation of the characteristic over time. The proportion of the sample matched over time has been determined from the design of the survey in terms of "designated" housing units. The actual proportion of the sample matched from year to year for any quarter for households and particularly for persons and families may be somewhat less than that for housing units due to sample attrition resulting from movements of people and changes in families. However, the impact on the efficiency of a composite estimator of level is generally likely to be negligible due to minor changes in the proportion of sample overlap. This is due to the fact that for estimates of levels, the optimum percent to be matched is no more than 50% (See Cochran p. 347), and the SIPP design provides 50% overlap for annual estimates and 50% or more overlap for quarterly estimates one year apart (See attachments II and III). Actual correlations for SIPP data are not yet available and we do not know how the correlation over time for a characteristic, say income, may differ for households, persons and families. A study of correlations of SIPP variables over time, and sample attrition and proportion of households, persons and families matched over time will have to be done later when relevant data are available. Consequently, our analysis of composite estimators for annual as well as quarterly estimates are given in terms of housing units matched over time and assumed correlations. Note that the variances of composite estimators given in this report will be applicable to household, person and family estimates when read with appropriate correlations obtained from SIPP data.

2. Annual Estimates

Composite estimates of annual levels and year to year change in levels can be obtained either from quarterly composite estimates or from simple annual estimates. First, we discuss the second procedure where three possible composite estimators are available to improve the reliability of annual estimates. These are given by Cochran (1963), Ernst-Breau (1983), and Wolter (1979). A brief description of these estimators are given in Attachment I. Relative variances of these estimators compared to the usual ratio estimator are given in Table 1. It can be seen that the composite estimation will reduce the variance of the annual estimate from 1% to 39% when correlation ranges from .2 to .9. (These are assumed correlations, actual correlations from SIPP data are not yet available). The variance of the estimate of change will be reduced by 1% to 59% for correlations ranging from .2 to .9. As mentioned earlier, the correlation between family incomes reported to the IRS on two consecutive years was found to be .80 to .87. The correlation between SIPP income data may be somewhat less because of the possibility of a higher response error in SIPP compared to income tax returns. However, when the correlation is .8, the variance of annual income can be reduced by 25% and the variance of the year to year change in income can be reduced by 38% by composite estimation. Thus, substantial gains in reliability of estimates can be achieved by composite estimation. Also, note that the gains in reliability from the composite estimation are much greater in estimating change than current level.

All three estimators are minimum variance estimators and are asymptotically equally efficient. The variances for Cochran and Ernst-Breau estimators given in Table 1 are limiting variances. However, from the third or fourth year of composite estimation variances are about the same as the limiting variances. The results from Wolter's method are based on assuming that ten years' data will be utilized in estimation. We have obtained the coefficients of simple annual estimates and variance-covariance matrices for Wolter's method for assumed correlations .2 to .9 and the results for correlation .8 are given in Table 2. Wolter's method improves the reliability marginally at middle year(s), for example for 5th and 6th year in estimation based on ten years' data. Also, note that the reliability of estimates of change is better in Wolter's method because estimates of levels for earlier years are revised using all data up to and including the latest year. The other two methods assume that estimates of levels will not be revised, although it is possible to do so. In practice, the revision of estimates published earlier may not be desirable.

Like both the K and AK CPS composite estimators, Cochran and Ernst-Breau estimators require only storage of the previous year's composite estimate for computing the current year's estimate. Wolter's method, on the other hand, requires storage of simple estimates (two simple estimates from two panels in each year) of all previous years. This presents a difficult problem in terms of data storage for SIPP with numerous data items. Also, Wolter's method is much more computationally complex. For example, with ten years' data, this requires inversion of 20x20 matrices and becomes a computational burden with a large number of survey characteristics. Data storage requirements and computational burden involved in Wolter's method can be eased by restricting the use of data to fewer years (say 5 years) with some loss in reliability. We have not yet evaluated the loss in efficiency that may result in using 5 years' data compared to all ten years' data. However, since all three estimators are minimum variance estimators, the choice of an estimator must depend on the data storage requirements and computational ease and on that basis either the Cohcran or Ernst-Breau estimator may be used for composite annual estimates.

Note that the panels that do not provide simple annual estimates can not be used for composite estimation of annual levels and change. This means that about 9% to 20% monthly data cannot be utilized in composite estimation based on simple annual estimates (see Attachment II). On the other hand, annual estimates obtained by pooling quarterly composite estimates will utilize almost all available monthly data and, therefore, likely to be much more efficient. Thus, alternative annual estimators based on pooling quarterly composite estimates must be researched and the discussion of this is deferred to Section 5.

3. Quarterly Estimates

It should be noted that the data available for estimation varies from quarter to quarter (see Attachment III). For analytical purposes, a panel that provides 12 observations (4 observations for each of the three months of a quarter) is considered 1 (full) panel. The panels providing less than 12 observations are expressed as fractional panels. For example, a panel providing 9 observations is a 3/4 panel. On this basis, data from 3, 2.75, 2 and 2.5 panels are available for 1st, 2nd, 3rd, and 4th quarters respectively (See attachment III). Consequently, the reliability of quarterly estimates will vary from quarter to quarter for the usual ratio (uncomposited) estimator as well as for any composite estimator. For a given sample size and for a given correlation of a characteristic over time, the efficiency of a composite estimator depends on proportions of the sample matched with earlier years. For quarters one year apart (for example, 1st quarters of 1986 and 1987), the proportions of sample matched one year apart are .6667, .6364, .5000 and .6000 for 1st, 2nd, 3rd, and 4th quarters respectively. And the proportions matched two years apart are .3333, .2727, 0.0, and .2000 for 1st, 2nd, 3rd, and 4th quarters respectively. Given this data structure, even if the same composite estimator is best for all quarters it may not be identical in the sense that the optimum coefficient K (See attachment I) may be different for different quarters. Note also, that composite estimation can be obtained beginning with the year 1987 for the 1st and 2nd quarters and beginning with the year 1986 for the 3rd and 4th quarters (See attachment III).

Wolter's method which is applicable to any data structure, can be used for all 4 quarters. Cochran's method can be used for all quarters if we consider only matching of the sample one year apart and ignore the fractional matching two years apart. However, Cochran's method can be generalized to take advantage of the sample matching two years apart. Ernst-Breau's method is applicable only for the 3rd quarter which has 50% sample overlap from year to year, but may be generalized for other quarters. All these three estimators are minimum variance estimators. The K composite estimator with K=.5 (See Attachment I) previously used in CPS also can be used but it will be generally less efficient than these three minimum variance estimators.

The relative variance of composite estimators of level and change are given in Tables 3, 4, 5, and 6 for four quarters. For comparison purposes, the variance computations are based only on samples one year apart although Wolter's method can use samples two years apart. From Table 3 we see that for the 1st quarter, the variance of the estimate of level can be reduced by 1% to 37%

when correlation ranges from .2 to .9 by Cochran's method. For the estimate of change in quarterly levels one year apart the reduction in variance is from 1% to 55%. The variances for Cochran's method are limiting variances that can be generally achieved by 3rd or 4th year of composite estimation. Variance reductions in earlier years are marginally less, for example, for $\rho = .7$ relative variances for level are .885, .857, .849 for the 1st, 2nd and 3rd years and the limiting variance .847 is obtained in the 4th year of composite estimation. Results for Wolter's method are given assuming that only two years' data will be used in estimation. Greater reductions in variances could be achieved with increasing years if estimates of previous years are revised using data for later years. For the 2nd quarter, the reduction in variance is from 1% to 38% for estimates of level and is from 1% to 56% for estimates of change when correlation ranges from .2 to .9 (Table 4). For the 3rd quarter, the reduction in variance is from 1% to 39% for estimates of level and is from 1% to 59% for estimates of change (Table 5). Finally, for the 4th quarter, the reduction in variance is from 1% to 39% for estimates of level and is from 1% to 57% for estimates of change (Table 6).

Note that for a given correlation, the reduction in variance by Cochran's method is of the same order for all four quarters. This is due to the fact that the variance reduction for the composite estimation of level is maximum when sample overlap is 50% and all quarters have 50% or more overlap from year to year. The optimum value of the coefficient K depends on the proportion of sample overlap and on the value of correlation.

These results suggests the possibility of using a 'compromise' value of K for all quarters and for all characteristics if correlations do not differ much for different characteristics. Thus, computation of correlations for SIPP characteristics and a study of the loss in efficiency of the estimator with a 'compromise' K instead of the optimal K is important. In practice, it may be desirable to use the same value of K for all characteristics so that the estimates in a table will retain simple additive property. Also, this makes the processing and tabulation of data simpler. On the other hand, if correlations differ considerably for different characteristics like income and person/households receiving certain benefits, use of different values of K for such characteristics will provide maximum benefits in terms of variance reduction. If SIPP characteristics can be grouped into two or three groups on the basis of correlation and if these groups do not appear in the same table in the published reports, then the use of optimal values of K for different groups of characteristics should be considered.

Because of the data storage requirements and computational complexities (for example, inversion of 30x30 matrices for 1st quarter estimates with ten years' data) involved in Wolter's method, Cochran's method appears to be preferable for estimation of quarterly levels and changes. Our research shows that Cochran's estimator is the minimum variance estimator if the estimation is based on samples one year apart only. Similarly, Cochran's method, when generalized to take advantage of the sample matching two years apart, should be the minimum variance estimator.

4. Correlation and Robustness of Estimators

As mentioned in the previous section, the optimal coefficients for Cochran's or Ernst-Breau's estimators will be different for different characteristics. In practice, however, the use of the same coefficient (i.e., same estimator) for all characteristics makes processing and tabulation of data simpler and estimates in a table retain simple additive property. Thus, a comprehensive study of correlations for various characteristics from SIPP data will be needed to determine a 'compromise' composite estimator to be used for quarterly estimates. Also, a study of robustness of an estimator with respect to departures from the actual correlation needs to be done to determine optimal coefficients for estimators. Ernst-Breau estimator appears to be reasonably robust for slight departures from the actual correlation. Robustness of the K composite estimator previously used in CPS needs also to be studied and compared with these estimators. Robustness will be a key factor in the choice of a 'compormise' composite estimator for quarterly estimates as well as for the annual estimate.

5. Consistency Between Quarterly and Annual Estimates

An annual estimate is consistent with quarterly estimates if it is the sum (or simple average) of four quarterly estimates.

As mentioned earlier, annual estimates can be obtained by two different methods. (1) Composite estimates based on simple annual estimates (discussed in Section 2) and (2) composite estimates based on quarterly composite estimates.

Composite estimates for annual levels and change based on simple annual estimates will of course not be consistent with quarterly estimates. This method is not likely to be used anyway because it cannot utilize all available data.

Annual estimates (monthly averages for the year) obtained by averaging four quarterly composite estimates will of course satisfy consistency requirements. However, the most efficient annual estimate that can be formed from quarterly composite estimates is the weighted average of four estimates; weights depend on the variance - covariance matrix of four estimates. Such an annual estimate, however, will not be consistent with quarterly estimates if variances are unequal. Efficiency of the consistent annual estimator can be compared with this estimator and the loss in efficiency for using the consistent estimator may be determined when actual correlations are available.

How important is the consistency of quarterly and annual estimates should be considered carefully in terms of data publication plans and general perception problems. If consistency is important, how much loss in reliability of annual estimates are we willing to accept to maintain consistency?

6. Other Approaches

It has been suggested that composite estimation be done separately for each month and then summed to obtain estimates for quarters, years or any other combinations of months. There is, however, a question whether such an approach will yield estimators that are reasonably close to optimal for quarterly and annual estimates. The correlation of quarterly incomes (monthly averages for a quarter) from year to year is likely to be higher and more stable than the correlation of monthly incomes from year to year. Thus, a composite quarterly estimate will be generally more effective in reducing the variance than a composite monthly estimate. The sample overlap for months one year apart varies more than the overlap for quarters one year apart. Consequently, a 'compromise' composite estimator for 12 months is likely to be less optimal than a 'compromise' quarterly estimator for 4 quarters. Also, a quarterly estimate consistent with monthly estimates is the one which is a simple average (or sum) of 3 monthly estimates. The most efficient quarterly estimate that can be obtained from monthly composite estimates is the weighted average of 3 monthly estimates. Such an estimate, however, will not be consistent with monthly estimates. Also, even this weighted average may not be as efficient as the minimum variance composite estimators for quarterly levels discussed in Section 3.

7. Time-in-Sample Bias

The existence of bias due to rotation groups (time-in-sample) has been well documented for CPS [Bailar 1975]. Similarly, time-in-sample bias is very likely to exist in the SIPP and this presents a problem much more complex than in CPS because time-in-sample varies considerably from month to month and quarter to quarter. For example, 1986 first quarter will have waves 8 and 9 from the 1984 Panel, 4 and 5 from the 1985 Panel and 1 and 5 from the 1986 Panel. Where as 1987 first quarter will use data from waves 7 and 8 from the 1985 Panel, 4 and 5 from the 1986 Panel and 1 and 2 from the 1987 Panel (See Attachment III). Comprehensive research is needed (a) to determine the magnitude of time-in-sample bias for key SIPP variables and (b) to develop procedures for possible elimination or at least reduction of time-in-sample bias.

8. Within Interview Correlated Response Bias

Unlike CPS, estimates from SIPP may include within interview bias resulting from the fact that people may tend to report similar amounts of income received for the four reference months in the same interview. If this is the case, the between-interview variability is expected to be greater than the within-interview variability for two consecutive months. This potential problem in association with composite estimation needs to be researched.

9. Composite Estimation for Longitudinal Reports

For longitudinal studies where we have eight repeated interviews for most of the households/persons for 2½ years, and the objective is to characterize patterns of individual response and change over time, time series methods based on auto-regressive model (Box and Jenkins 1970) may be most suitable. Other methods such as general multivariate and random effects models (Rao '67 '75) may also be useful.

Estimates like monthly average number of persons/households receiving certain benefits, for example, food stamps, in a quarter and in a year do not reveal all information. One might also be interested in the distribution of persons/households receiving certain benefits by duration of the receipt of benefits. Such analysis of data can be made from longitudinal files. Research on this topic is to be persued as a separate project as longitudinal data files are developed.

10. Integration of Cross-Sectional and Longitudinal Estimation

The extend to which cross-sectional and longitudinal composite estimation can be integrated should be researched after optimal procedures have been developed for each independently. Consequently, research on this topic will be taken up later.

Table 1. The Relative Variance of Composite Estimators for Annual Levels and Change

Correlation	Composite Estimators		Relative Variance of Level	Relative Variance of Change
.2	Cochran	K = .4949 [3]	.990	.989
	Wolter*		.990[1]	.983[1]
			.990[2]	.977[2]
.3	Cochran	K = .4882 [3]	.976	.972
	Wolter*		.976[1]	.959[1]
			.954[2]	.950[2]
.4	Cochran	K = .4782 [3]	.956	.946
	Wolter*		.956[1]	.922[1]
			.905[2]	.907[2]
.5	Cochran	K = .4641[3]	.928	.906
	Wolter*		.928[1]	.867[1]
			.866[2]	.845[2]
.6	Cochran	K = .4444[3]	.889	.847
	Wolter*		.889[1]	.790[1]
			.800[2]	.762[2]
.7	Cochran	K = .4166[3]	.833	.758
	Wolter*		.833[1]	.682[1]
			.714[2]	.650[2]
.8	Cochran	K = .3750[3]	.750	.625
	Wolter*		.750[1]	.531[1]
			.600[2]	.500[2]
.9	Cochran	K = .3036[3]	.607	.412
	Wolter*		.607[1]	.317[1]
			.456[2]	.297[2]

Note: The proportion of the sample (housing units) matched between two consecutive years is .5. The results from the Ernst-Breau Estimator are the same as those from the Cochran Estimator.

3/ Optimum value of K.
* Assumes ten years' data will be used and estimates for earlier years will be revised using later years' data.
1/ Relative variance for end (1, 10) years.
2/ Relative variance for middle (5th and 6th) years.

Table 2. Optimum Coefficients for Wolter's Estimators[19] of Annual Levels ($\rho = .8$)

										Estimator
.3750004	.6249996	-.3124989	.3124989	-.1562477	.1562477	-.0781203	.0781203	-.0390530	.0390530	$\hat{X}_{10}$
-.0195122	.0195122	-.0097275	.0097275	-.0048065	.0048065	-.0022888	.0022888	-.0009155	.0009155	
.1875008	-.1875008	.4687517	.5312481	-.2656210	.2656210	-.1328044	.1328044	-.0663901	.0663901	$\hat{X}_9$
-.0331707	.0331707	-.0165367	.0165367	-.0081711	.0081711	-.0038910	.0038910	-.0015564	.0015564	
.0937515	-.0937515	.2343786	-.2343786	.4921951	.5078049	-.2538908	.2538908	-.1269222	.1269222	$\hat{X}_8$
-.0634146	.0634146	-.0316143	.0316143	-.0156212	.0156212	-.0074387	.0074387	-.0029755	.0029755	
.0468779	-.0468779	.1171947	-.1171947	.2461088	-.2461088	.4980774	.5019226	-.2509153	.2509153	$\hat{X}_7$
-.1253658	.1253658	-.0624991	.0624991	-.0308819	.0308819	-.0147057	.0147057	-.0058823	.0058823	
.0234432	-.0234432	.0586081	-.0586081	.1230769	-.1230769	.2490843	-.2490843	.4996338	.5003662	$\hat{X}_6$
-.2499998	.2499998	-.1246333	.1246333	-.0615835	.0615835	-.0293255	.0293255	-.0117302	.0117302	
.0117302	-.0117302	.0293255	-.0293255	.0615835	-.0615835	.1246333	-.1246333	.2499998	-.2499998	$\hat{X}_5$
.5003662	.4996338	-.2490843	.2490843	-.1230769	.1230769	-.0586081	.0586081	-.0234432	.0234432	
.0058823	-.0058823	.0147057	-.0147057	.0308819	-.0308819	.0624991	-.0624991	.1253658	-.1253658	$\hat{X}_4$
.2509153	-.2509153	.5019226	.4980774	-.2461088	.2461088	-.1171947	.1171947	-.0468779	.0468779	
.0029755	-.0029755	.0074387	-.0074387	.0156212	-.0156212	.0316143	-.0316143	.0634146	-.0634146	$\hat{X}_3$
.1269222	-.1269222	.2538908	-.2538908	.5078049	.4921951	-.2343786	.2343786	-.0937515	.0937515	
.0015564	-.0015564	.0038910	-.0038910	.0081711	-.0081711	.0165367	-.0165367	.0331707	-.0331707	$\hat{X}_2$
.0663901	-.0663901	.1328044	-.1328044	.2656210	-.2656210	.5312481	.4687519	-.1875008	.1875008	
.0009155	-.0009155	.0022888	-.0022888	.0048065	-.0048065	.0097275	-.0097275	.0195122	-.0195122	$\hat{X}_1$
.0390530	-.0390530	.0781203	-.0781203	.1562477	-.1562477	.3124989	-.3124989	.6249996	.3750004	

Table 2a. Variance - Covariance Matrix of Annual Estimates
(Relative Variance - Covariance)

.37500045	.18750076	.09375145	.04687787	.02344323	.01173020	.00588226	.00297546	.00155640	.00091553
.18750076	.31875129	.15937747	.07969238	.03985349	.01994133	.00999985	.00505827	.00264507	.00155640
.09375145	.15937747	.30469222	.15235308	.07619049	.03812314	.01911736	.00967026	.00505829	.00297546
.04687787	.07969238	.15235308	.30119033	.15062274	.07536652	.03779355	.01911736	.00999985	.00588226
.02344323	.03985349	.07619049	.15062274	.30036635	.15029315	.07536652	.03812314	.01994133	.01173020
.01173020	.01994133	.03812314	.07536652	.15029315	.30036635	.15062274	.07619049	.03985349	.02344323
.00588226	.00999985	.01911736	.03779355	.07536652	.15062274	.30119033	.15235308	.07969238	.04687787
.00297546	.00505829	.00967026	.01911736	.03812314	.07619049	.15235308	.30469222	.15937747	.09375145
.00155640	.00264507	.00505829	.00999985	.01994133	.03985349	.07969238	.15937747	.31875129	.18750076
.00091553	.00155640	.00297546	.00588226	.01173020	.02344323	.04687787	.09375145	.18750076	.37500045

$\times\ 2$

Table 3. The Relative Variance of Composite Estimators for Quarterly Levels and Change, First Quarter

First Quarter

$\lambda = 2/3$[1]

Correlation ρ	Composite Estimators		Relative Variance of Level	Change
.2	Cochran	K = .3303	.991	.990
	Wolter*		.991, .991	.989
.3	Cochran	k = .3263	.979	.976
	Wolter*		.980, .980	.972
.4	Cochran	k - .3203	.961	.954
	Wolter*		.964, .964	.944
.5	Cochran	K = .3117	.935	.920
	Wolter*		.943, .943	.900
.6	Cochran	K = .2997	.899	.869
	Wolter*		.917, .917	.833
.7	Cochran	K = .2823	.847	.790
	Wolter*		.885, .885	.734
.8	Cochran	K = .2558	.767	.665
	Wolter*		.847, .847	.584
.9	Cochran	K = .2091	.628	.452
	Wolter*		.802, .802	.358

* Based on two years 1986-1987. Two variances for level are for '86 and '87.

1 λ is the proportion of the sample (housing units) matched between two years for the quarter

Table 4. The Relative Variance of Composite Estimators for Quarterly Levels and Change, Second Quarter

Second Quarter

$\lambda = .6364$[1]

Correlation ρ	Composite Estimators		Relative Variance of Level	Relative Variance of Change
.2	Cochran	K = .3602	.990	.990
	Wolter*		.990, .993	.990
.3	Cochran	K = .3556	.978	.975
	Wolter*		.978, .984	.975
.4	Cochran	K = .3489	.959	.952
	Wolter*		.962, .972	.951
.5	Cochran	K = .3391	.933	.916
	Wolter*		.940, .957	.913
.6	Cochran	K = .3256	.896	.863
	Wolter*		.913, .936	.856
.7	Cochran	K = .3063	.842	.781
	Wolter*		.880, .911	.772
.8	Cochran	K = .2769	.761	.654
	Wolter*		.841, .878	.643
.9	Cochran	K = .2256	.620	.440
	Wolter*		.794, .832	.434

* Based on two years 1986-1987. Two variances for level are for '86 and '87.

1 λ is the proportion of the sample (housing units) matched between two years for the quarter.

Table 5. The Relative Variance of Composite Estimators for Quarterly Levels and Change, Third Quarter

Third Quarter

λ = 1/2[1]

Correlation ρ	Composite Estimators		Relative Variance of Level	Relative Variance of Change
.2	Cochran	K = .4949	.990	.989
	Wolter*		.990, .990	.988
.3	Cochran	K = .4882	.976	.972
	Wolter*		.977, .977	.969
.4	Cochran	K = .4782	.956	.946
	Wolter*		.958, .958	.937
.5	Cochran	K = .4641	.928	.906
	Wolter*		.933, .933	.889
.6	Cochran	K = .4444	.889	.847
	Wolter*		.901, .901	.816
.7	Cochran	K = .4166	.833	.758
	Wolter*		.860, .860	.710
.8	Cochran	K = .3750	.750	.625
	Wolter*		.810, .810	.565
.9	Cochran	K = .3036	.607	.412
	Wolter*		.746, .746	.331

Note: Results for the Ernst-Breau estimator which is applilcable because λ = .5 , are the same as those from the Cochran Estimator.

* Based on two years, 1986-1987. Two variances for level are for '86 and '87.

1 λ is the proportion of sample matched between two years for the quarter.

Table 6. The Relative Variance of Composite Estimators for Quarterly Levels and Change, Fourth Quarter

Fourth Quarter

λ = .6000[1]

Correlatoin ρ	Composite Estimators		Relative Variance of Level	Relative Variance of Change
.2	Cochran	K = .3961	.990	.989
	Wolter*		.994, .990	.991
.3	Cochran	K = .3909	.977	.974
	Wolter*		.987, .978	.977
.4	Cochran	K = .3832	.958	.950
	Wolter*		.976, .961	.954
.5	Cochran	K = .3723	.931	.912
	Wolter*		.962, .938	.918
.6	Cochran	K = .3570	.892	.856
	Wolter*		.942, .910	.863
.7	Cochran	K = .3352	.838	.772
	Wolter*		.915, .876	.778
.8	Cochran	K = .3024	.756	.643
	Wolter*		.877, .834	.647
.9	Cochran	K = .2456	.614	.429
	Wolter*		.820, .782	.426

* Based on two years 1986-1987. Two variances for level are for '86 and '87.

1 λ is the proportion of the sample (housing units) matched between two years for the quarter.

11. Summary

A. General Estimates

1. Estimates of quarterly levels (monthly averages for quarters), annual levels and year to year change in quarterly and annual levels are of major interest for cross-sectional reports.
2. SIPP rotation pattern creates a data structure that varies from month to month and quarter to quarter.
3. The AK Composite Estimator currently used in CPS is appropriate for CPS type (4-8-4 or 3-9-3) rotation patterns but cannot be used for SIPP data structure.
4. The K Composite Estimator previously used for CPS can be used but other estimators will be more efficient.

B. Quarterly Estimates

1. Data available for estimation varies from quarter to quarter as shown below.

Quarter	No. of Panels [1]	Proportion of Sample Matched	
		One Year Apart	Two Years Apart
1	3	.6667	.3333
2	2.75	.6364	.2727
3	2	.5000	0.0
4	2.5	.6000	.2000

[1] A panel that provides 12 observations, 4 for each of the three months of a quarter, is considered 1 (full) panel.

2. Reliability of the usual ratio estimator as well as of any composite estimator will vary from quarter to quarter.

3. Composite Estimation

Three minimum variance composite estimators are available for quarterly estimates. Their applicability, advantages and disadvantages are as follows:

Minimum Variance Composite Estimators	Applicability and Advantages	Disadvantages
Wolter (1979)	Applicable to all 4 quarters.	Requires storage of simple estiates for all previous years. Computationally complex.
Cochran (1963)	Applicable to all 4 quarters. Requires storage of composite estimate for the previous year only.	Ignores sample overlap two years apart but may be generalized to include this.
Ernst-Breau (1983)	Applicable to 3rd quarter only. Requires storage of composite estimate for the previous year only.	Not applicable to 1st, 2nd & 4th quarter. Ignores sample overlap two years apart. Needs to be generalized for such applications.

4. Expected Variance Reduction by Composite Estimation

Reduction in Variance*
(Percent)

Quarter	Level Correlation								Change Correlation							
	.2	.3	.4	.5	.6	.7	.8	.9	.2	.3	.4	.5	.6	.7	.8	.9
1	1	2	4	6	10	15	23	37	1	2	5	8	13	21	34	55
2	1	2	4	7	10	16	24	38	1	2	5	8	14	22	35	56
3	1	2	4	7	11	17	25	39	1	3	5	9	15	24	37	59
4	1	2	4	7	11	16	24	39	1	3	5	9	14	23	36	57

* Based on Cochran's method which ignores sample overlap two years apart. Reductions will be somewhat higher if the procedure is generalized to include fractional overlaps two years apart.

C. Annual Estimates and Consistency Between Quarterly and Annual Estimates

1. Composite estimation based on "Simple Annual Estimates".

   Three minimum variance estimators (see 11.B) are applicable but this method is not optimal since it does not utilize all available data.

2. Annual estimates based on composite quarterly estimates will be more efficient than those based on "Simple Annual Estimates".

3. A consistent annual estimate that is a simple mean of composite quarterly estimates will not be the most efficient estimate for annual levels.

D. Further Research Needed

1. Compute correlations over time for various SIPP characteristics.

2. Study robustness of composite estimators with respect to departures from the actual correlation.

3. Determine the "compromise coefficient" to be used for Cochran/Ernst-Breau estimators based on correlation and robustness analysis.

4. Consistency of annual and quarterly estimates

   a. Determine the importance of consistency.

   b. Develop the most efficient composite estimator for annual levels and change.

   c. Evaluate the loss in efficiency for using a consistent annual estimator instead of the most efficient annual estimator.

   d. Determine the composite estimator to be used for annual estimates weighing the importance of consistency against the loss in reliability.

5. Generalize Cochran's method to take advantage of sample overlap two years apart.

6. Generalize Ernst-Breau's method for all quarters if possible.

7. Determine whether Wolter's method provides 'unique' solution when sample overlap two years apart are included in the model.

8. Evaluate the loss in efficiency in using Wolter's method for shorter period (say 5 years) compared to 10 years.

9. Time-in-Sample Bias.

   Evaluate the magnitude of time-in-sample bias and develop procedures for possible elimination or at least reduction of time-in-sample bias.

10. Within Interview Correlated Response Bias.

    This potential problem in association with composite estimation needs to be researched.

11. Integration of Cross-Sectional and Longitudinal Estimates.

    Determine the extent to which cross-sectional and longitudinal estimates can be integrated.

References

1. Bailar, B. A. (1975), "The Effects of Rotation Group Bias on Estimates from Panel Surveys". JASA, 70, 23-29.

2. Box, G.E.P., and Jenkins, G.M. (1970), "Time Series Analysis, Forecasting and Control", Holden day, San Francisco.

3. Cochran, W.G. (1963), "Sampling Techniques" John Wiley and Sons, New York.

4. Ernst, L., and Breau, P. (1983), unpublished manuscript.

5. Gurney, M. and Daly, J.F. (1965), "A Multivariate Approach to Estimation in Periodic Surveys," Proceedings of the American Statistical Association, Social Statistics Section, 242-257.

6. Hansen, M.H., Hurwitz, W.N. and Madow, W.G. (1953), "Sample Survey Methods and Theory" Vol. II. John Wiley & Sons, New York.

7. Huang, E.T. and Ernst, L.R. (1981), "Comparison of An Alternative Estimator to the Current Composite Estimator in CPS", Proceedings of the American Statistical Association, Survey Research Methods Section 303-308.

8. Nelson, D., McMillen, D. and Kasprzyk, D. (1984), "An Overview of the Survey of Income and Program Participation", SIPP Working Paper Series No. 8401. U.S. Dept. of Commerce, Bureau of the Census, Washington, D.C.

9. Ponikowski, C. and Tadros, W. (1984), "Correlation Between Year to Year Incomes" Census Bureau Memorandum for Documentation, May 9, 1984.

10. Rao, C.R. (1965), "The Theory of Least Squares When Parameters are Stochastic and Its Application to Analysis of Growth Curves", Biometrika 52, 447-458.

11. __________1975), "Simulataneous Estimation of Parameters of Different Linear Models and Application to Biometric Problems", Biometrics, 31, 545-554.

12. U.S. Dept of Commerce, Bureau of the Census, (1978), "The Current Population Survey: Design & Methodology" Technical Paper 40. Washington, D.C.

13. Wolter, K.M. (1979), "Composite Estimation in Finite Populations", JASA, 74, 604-613.

Attachment I

Composite Estimation in the CPS

1. The K Composite Estimator

"The Current Population Survey - Design and Methodology," Technical Paper 40 (1978).

The K composite estimator of level (mean or total of a certain labor force characteristic) in the current month h, y'_h is given by

$$y'_h = (1-K)\, y_h + K\,(y'_{h-1} + d_{h,\, h-1});\ 0 < K < 1$$

where

$$d_{h,\, h-1} = y''_h - y''_{h-1}$$

y'_{h-1} is the composite estimator of level for the preceding month h-1

y_h is the second-stage ratio estimator of level for the current month h

y''_h and y''_{h-1} are the second-stage ratio estimators for months h and h-1 based on the six rotation groups that are in sample for both months h and h-1.

The variance of this estimator is given in Hansen, Hurwitz and Madow (1953). This estimator with K = .5 was used previously in the CPS.

2. The AK Composite Estimator

Huang and Ernst (1981) "Comparison of an Alternative Estimator to the Current Composite Estimator in CPS."

Under the 4-8-4 rotation pattern, let y_{hi} be the second stage ratio estimator of the level (mean or total) in the current month h based on the rotation group i (which is in its i-th time in the sample), $i = 1, 2, \ldots 8$. Assume the following:

(1) $V(y_{hi}) = \sigma^2$ for all h, i

(2) Estimators derived from different rotation groups of a given month are uncorrelated; that is

$Cov\ (y_{hi}, y_{hj}) = 0\ ,\ i \neq j = 1, 2, \ldots 8$

(3) Estimators derived from the overlapping rotation groups are covariance statioinary.

The AK composite estimator, first defined by Gurney and Daly (1965), is a generalization of the K composite estimator previously used in CPS and is given by

$$y'_h = (1/8)\{(1-K+A)\ (y_{h1} + y_{h5}) + (1-K-A/3)$$
$$(y_{h2} + y_{h3} + y_{h4} + y_{h6} + y_{h7} + y_{h8})\}$$
$$+ K\ (y'_{h-1} + d_{h,h-1})\ ,\ 0 \leq A \leq 1,\ 0 \leq K \leq 1$$

where
$$d_{h,h-1} = (1/6)\ (y_{h2} + y_{h3} + y_{h4} + y_{h6} + y_{h7} + y_{h8}$$
$$-(\ y_{h-1,\ 1} + y_{h-1,\ 2} + y_{h-1,\ 3} + y_{h-1,5}$$
$$+ y_{h-1,\ 6} + y_{h-1,\ 7}))$$

Note that when $A = 0$, $K = 0$ the AK composite estimator reduces to the simple average of the estimates from the eight rotation groups. When $A = 0$ and $K = .5$ the AK composite estimator reduces to the K composite estimator previously used in the CPS.

Thus, the AK composite estimator has the potential of assigning more weight to the rotation groups which have been in the sample for the 1st and 5th time, and less weight to the rest of the rotation groups than the corresponding weights in the K composite estimator.

The variance of the AK composite estimator is given in Huang and Ernst (1981). The AK composite estimator was found to be more efficient than the K composite estimator for monthly level, month to month change and annual average for both 4-8-4 and 3-9-3 rotation patterns. The AK composite estimator with A = .2 and K = .4 is currently being used in the CPS.

Minimum Variance Composite Estimators

1. Wolter (1979) "Composite Estimation in Finite Populations", JASA.

This procedure is applicable to any rotation pattern. Illustrated here for first quarter estimates one year apart considering only two years' data.

For 1987 and 1986 1st quarters, the vector of simple estimators is

$$Y = (y_{t\ t},\ y_{t\ t-1}, y_{t\ t-2}, y_{t-1\ t-1}, y_{t-1\ t-2}, y_{t-1\ t-3})'$$

where y_{ij} is the first quarter simple estimate (monthly average) for the i-th year from the j-th year panel. (t = 87, t-1 = 86. etc.)

Let $\beta = (x_t, x_{t-1})'$

be the vector of unknown parameters. (true monthly averages for the quarter).

Assume the general linear model

$$Y = X\beta + e$$

the error vector e satisfies

$$E(e) = 0,\ E(e\,e') = V$$

V is the variance - covariance matrix of the vector of simple estimators and must be nonsingular. We have verified algebraically that V is nonsingular for annual estimates, and for all quarterly estimates considering sample overlap one year apart only.

Then

$$\hat{\beta} = PY$$

where

$$P = (X' V^{-1} X)^{-1} X' V^{-1}$$, V^{-1} is the inverse of matrix V

and $$Var(\hat{\beta}) = (X' V^{-1} X)^{-1} = C$$ say.

The design matrix X and variance - covariance matrix V in this case are

$$X = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \qquad V = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \rho & 0 & 0 \\ 0 & 0 & 1 & 0 & \rho & 0 \\ 0 & \rho & 0 & 1 & 0 & 0 \\ 0 & 0 & \rho & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Where $\sigma^2 = 1$ (assumed) and ρ is the correlation coefficient.

The matrix P is composed of the coefficients that are applied to the simple estimators in forming the minimum variance linear unbiased (MVLU) estimators. Thus,

$$\hat{X}_t = P_1 Y , \quad \hat{X}_{t-1} = P_2 Y$$

where P_i is the i-th row of P.

The estimator of change $X_t - X_{t-1}$ is simply $\hat{X}_t - \hat{X}_{t-1} = P_1 Y - P_2 Y$

$$Var(\hat{X}_t - \hat{X}_{t-1}) = C_{11} + C_{22} - 2C_{12}$$

where C_{ij} is the ij element of C matrix.

It is important to note that this estimator uses all available data up to and including the year t. When data from succeeding years 88, 89 etc. are available, estimates for 87, 86 etc. also change.

Advantage: Applicable to any rotation pattern and data structure. It requires only specifications for design and variance-covariance matrices X and V.

Disadvantages:

(1) Requires storage of simple estimates for all previous years.

(2) Composite estimates for earlier years are changed.

(3) Becomes a computational burden with large number of survey characteristics as number of years increases because it requires inversion of matrices V and $(X' V^{-1} X)$.

2. Cochran (1963) Sampling Techniques

Assume sample size on each occasion is n. On occasion h, the proportion of the sample matched with occasion (h-1) is λ . The variance σ^2 and the correlation ρ between item values on the same units on two successive occasions are assumed same throughout. A part of the sample on occasion h may also be matched with the sample on occasion (h-2), but this is ignored. The two estimators of mean $\bar{Y}_h$ on the h-th occasion are $\bar{y}_{hu}$ from the unmatched proportion of the sample, and $\bar{y}'_{hm} = \bar{y}_{hm} + b\,(\bar{y}'_{h-1} - \bar{y}_{h-1,m})$ from the matched portion of the sample where b is the regression coefficient and $\bar{y}'_{h-1}$ is the composite estimator on occasion h-1. Then the composite estimator for the h-th occasion is

$$\bar{y}_h' = K\,\bar{y}_{hu} + (1-K)\,\bar{y}_{hm}'$$

and its variance is

$$V(\bar{y}_h') = \frac{\sigma^2}{n}\,g_h$$

where

$$g_h = \frac{K^2}{1-\lambda} + \frac{(1-K)^2(1-\rho^2)}{\lambda} + \rho^2(1-K)^2\,g_{h-1}$$

The asymptotic value of $V(\bar{y}_h')$ as $h \to \infty$ is

$$V(\bar{y}_\infty') = \frac{\sigma^2}{n}\,g_\infty\ ;$$

where

$$g_\infty = \frac{\lambda K^2 + (1-\lambda)(1-K)^2(1-\rho^2)}{\lambda(1-\lambda)[1-\rho^2(1-K)^2]} .$$

The value of K which minimizes the limiting variance $V(\bar{y}_\infty')$ is

$$K_{opt} = \frac{\sqrt{(1-\rho^2)}\,[\sqrt{(1-\rho^2) + 4\,\lambda(1-\lambda)\rho^2)} - \sqrt{(1-\rho^2)}]}{2\lambda\rho^2}$$

The estimator of change $\bar{Y}_h - \bar{Y}_{h-1}$ is simply $\bar{y}_h' - \bar{y}_{h-1}'$

And its variance is

$$V\,(\bar{y}_h' - \bar{y}_{h-1}') = \frac{\sigma^2}{n}\,[g_h + g_{h-1}\{1-2\rho(1-K)\}]$$

Advantages:

(1) Requires storage of the composite estimate for the previous year only.

(2) Composite estimates for earlier years are not changed.

(3) It is also the minimum variance estimator.

(4) Computationally much simpler than the Wolter estimator.

Disadvantage: This estimator ignores any overlap of the sample on the h-th occasion with (h-2) occasion. However, this procedure can be generalized to take advantage of such overlapping samples.

3. Ernst-Breau (1983). Unpubished Manuscript

This composite estimator was developed specifically for the CPS Supplements that are repeated every year like the March Supplement and thus assumed 50% overlap between samples one year apart and no overlap between samples two years apart.

The estimator of the current yearly level is

$$Y_1 = (1-A_1)\, X_{1,1} + A_1 X_{1,2} - A_2 X_{2,1} + A_2 X_{2,2} - A_3 X_{3,1} + A_3 X_{3,2} - \ldots\ldots\ldots\ldots\ldots - A_n X_{n,1} + A_n X_{n,2}$$

where, $X_{1,1}$ is the simple estimate of yearly level for the current year based on the 50% of the sample that is not matched with the previous year.

$X_{1,2}$ is the simple estimate of yearly level for the current year based on the 50% of the sample that is matched with the previous year.

and

$X_{2,1}$ is the simple estimate of yearly level for the previous year based on ther 50% of the sample that is matched with the current year, etc.

Assume $V(X_{i,j}) = \sigma^2$, for all i,j

$Cov\ (X_{i,2}, X_{i+1,1}) = \rho\sigma^2$ for all i

and all other covariances are zero.

Then coefficients A_i's are given by

$$A_i = C_1\ Z_1^{i-1} + C_2\ Z_2^{i-1}$$

where $Z_1 = [1 - \sqrt{(1-\rho^2)}]/\rho$

$$Z_2 = [1 + \sqrt{(1-\rho^2)}]/\rho$$

$$C_1 = (2 - \rho Z_1)/[(2 - \rho Z_1)^2 - Z_1^{2n-2}\ (2 - \rho Z_2)^2]$$

$$C_2 = -Z_1^{2n-2}\ (2 - \rho Z_2)/[(2 - \rho Z_1)^2 - Z_1^{2n-2}(2 - \rho Z_2)^2]$$

$$\lim_{n \to \infty} c_1 = 1\ /\ (2 - \rho Z_1)$$

$$\lim_{n \to \infty} C_2 = 0$$

Therefore, $\lim_{n \to \infty} Ai = Z_1^{i-1}\ /\ (2 - \rho Z_1)$

Finally, the limiting variance (algebraic details omitted) is

$$V(Y_1) = \sigma^2\ \{1 - \frac{2}{(2-\rho Z_1)} + \frac{2(1-\rho Z_1)}{(2-\rho Z_1)^2(1-Z_1{}^2)}\}$$

Advantages: (1) Requires storage of the composite estimate for the previous year only.

(2) Composite estimates for earlier years are not changed.

(3) It is also the minimum variance estimator.

(4) Computationally much simpler than the Wolter estimator.

Disadvantages: (1) This estimator can not be used for quarterly estimates except for the 3rd quarter which has 50% sample overlap from year to year.

(2) Also it ignores any overlap of the sample two years apart. This procedure may be generalized so that it can be applied for all quarters.

Attachment II.

Data Structure for Annual Estimates

Annual Estimate

1984 Estimate

1984 Panel

Wave	Rotation
2	1 2 3
3	4 1 2 3
4	4 1 2 3
5	4 1 2 3

1985 Panel

Wave	Rot	Data for
1	2	OND 1/
	3	ND
	4	D

(Insufficient) 2/

1985 Estimate

1984 Panel

Wave	Rotation
5	1 2 3
6	4 1 2 3
7	4 1 2 3
8	4 1 2 3

1985 Panel

Wave	Rot
1	2 3 4 1
2	2 3 4
3	1 2 3 4
4	1 2 3 4

1986 Panel

Wave	Rot	Data for
1	2	OND 1/
	3	ND
	4	D

(Insufficient) 2/

1986 Estimate

1984 Panel

Wave	Rotation	Data For
8	1	J 1/
	2	JF
	3	JFM
9	4	JFMA
	1	FMAM
	2	MAMJ
	3	AMJJ

(Insufficient) 2/

1985 Panel

Wave	Rot
4	2 3 4
5	1 2 3 4
6	1 2 3 4
7	1 2 3 4

1986 Panel

Wave	Rot
1	2 3 4 1
2	2 3 4
3	1 2 3 4
4	1 2 3 4

1987 Panel

Wave	Rot	Data for
1	2	OND 1/
	3	ND
	4	D

(Insufficient) 2/

1/ Letters J, F, ... D stand for months January, February ... December.

2/ This panel does not provide an annual estimate for the year e.g., 1985 Panel provides data only for the fourth quarter of 1984.

Attachment II (cont.)

Data Structure for Annual Estimates

1987 Estimate

1985 Panel

Wave	Rot	Data for
7	2	J [1]
	3	JF
	4	JFM
8	1	JFMA
	2	FMAM
	3	MAMJ
	4	AMJJ

(Insufficient) [2]

1986 Panel

Wave	Rot
4	2 3 4
5	1 2 3 4
6	1 2 3 4
7	1 2 3 4

1987 Panel

Wave	Rot
1	2 3 4 1
2	2 3 4
3	1 2 3 4
4	1 2 3 4

1988 Panel

Wave	Rot	Data for
1	2	OND [1]
	3	ND
	4	D

(Insufficient) [2]

1988 Estimate

1986 Panel

Wave	Rot	Data for
7	2	J [1]
	3	JF
	4	JFM
8	1	JFMA
	2	FMAM
	3	MAMJ
	4	AMJJ

(Insufficient) [2]

1987 Panel

Wave	Rot
4	2 3 4
5	1 2 3 4
6	1 2 3 4
7	1 2 3 4

1988 Panel

Wave	Rot
1	2 3 4 1
2	2 3 4
3	1 2 3 4
4	1 2 3 4

1989 Panel

Wave	Rot	Data for
1	2	OND [1]
	3	ND
	4	D

(Insufficient) [2]

[1] Letters J, F, ...D stand for months January, February, ...December.

[2] This panel does not provide an annual estimate for the year e.g., 1985 Panel provides data only for the two quarters of 1987.

Attachment III

Data Structure for Quarterly Estimates

Quarterly Estimates

First Quarter January, February, March

1984 Q_1

1984 Panel

W	R	Int M.	Data for
2	1	2/84	J
	2	3/84	JF
	3	4/84	JFM
3	4	5/84	JFM
	1	6/84	FM
	2	7/84	M

1985 Q_1

1984 Panel

W	R	Int M.	Data for
5	1	2/85	J
	2	3/85	JF
	3	4/85	JFM
6	4	5/85	JFM
	1	6/85	FM
	2	7/85	M

1985 Panel

W	R	Int M.	Data for
1	2	2/85	J
	3	3/85	JF
	4	4/85	JFM
	1	5/85	JFM
2	2	6/85	FM
	3	7/85	M

1986 Q_1

1984 Panel

W	R	Int M.	Data for
8	1	2/86	J
	2	3/86	JF
	3	4/86	JFM
9	4	5/86	JFM
	1	6/85	FM
	2	7/86	M

1985 Panel

W	R	Int M.	Data for
4	2	2/86	J
	3	3/86	JF
	4	4/86	JF
5	1	5/86	JFM
	2	6/86	FM
	3	7/86	M

1986 Panel

W	R	Int M.	Data for
1	2	2/86	J
	3	3/86	JF
	4	4/86	JFM
5	1	5/86	JFM
	2	6/86	FM
	3	7/86	M

1987 Q_1

1985 Panel

W	R	Int M.	Data for
7	2	2/87	J
	3	3/87	JF
	4	4/87	JFM
8	1	5/87	JFM
	2	6/87	FM
	3	7/87	M

1986 Panel

W	R	Int M.	Data for
4	2	2/87	J
	3	3/87	JF
	4	4/87	JFM
5	1	5/87	JFM
	2	6/87	FM
	3	7/87	M

1987 Panel

W	R	Int M.	Data for
1	2	2/87	J
	3	3/87	JF
	4	4/87	JFM
	1	5/87	JFM
2	2	6/87	JFM
	3	7/87	M

1987 Q_1 3 Panels

2 Panels matched with 1986 (λ_1 = .6667)

1 Panel matched with 1985 (λ_2 = .3333)

Legend: W = Wave, R = Rotation Group, Int. M. = Interview Month
Letters J, F,...D stand for months January, February, ...December.

Attachment III (cont.)

Second Quarter April, May, June

1984 Q_2				1985 Q_2				1986 Q_2				1987 Q_2			
1984 Panel				1984 Panel				1984 Panel							
W	R	Int M.	Data for	W	R	Int M.	Data for	W	R	Int M.	Data for	W	R	Int M.	Data for
3	4	5/84	A	6	4	5/85	A	9	4	5/86	A				
	1		AM		1		AM		1		AM				
	2		AMJ		2		AMJ		2		AMJ				
	3		AMJ		3		AMJ		3	8/86	AMJ				
4	4		MJ	7	4		MJ								
	1	10/84	J		1	10/85	J								
				1985 Panel				1985 Panel				1985 Panel			
				1	1	5/85	A	5	1	5/86	A	8	1	5/87	A
				2	2		AM		2		AM		2		AM
					3		AMJ		3		AMJ		3		AMJ
					4		AMJ		4		AMJ		4	8/87	AMJ
				3	1		MJ	6	1		MJ				
					2	10/85	J		2	10/86	J				
								1986 Panel				1986 Panel			
								1	1	5/86	A	5	1	5/87	A
								2	2		AM		2		AM
									3		AMJ		3		AMJ
									4		AMJ		4		AMJ
								3	1		MJ	6	1		MJ
									2	10/86	J		2	10/87	J
												1987 Panel			
												1	1	5/87	A
												2	2		AM
													3		AMJ
													4		AMJ
												3	1		MJ
													2	10/87	J

1987 Q_2 2.75 Panels

1.75 matched with 1986 (λ_1 = .6364)

.75 matched with 1985 (λ_2 = .2727)

A full panel has 12 observations; 4 observations for each of the three months of a quarter.

Attachment III (cont.)

Third Quarter July, August, September

1984 Q_3

1984 Panel

W	R	Int M.	Data for
3	3	8/84	J
4	4		JA
	1		JAS
	2		JAS
	3		AS
5	4	1/85	S

1985 Q_3

1984 Panel

W	R	Int M.	Data for
6	3	8/85	J
7	4		JA
	1		JAS
	2		JAS
	3		AS
8	4	1/86	S

1985 Panel

W	R	Int M.	Data for
2	4	8/85	J
3	1		JA
	2		JAS
	3		JAS
	4		AS
4	1	1/86	S

1986 Q_3

1984 Panel

W	R	Int M.	Data for
9	3	8/86 (omit)	J

1985 Panel

W	R	Int M.	Data for
5	4	8/86	J
6	1		JA
	2		JAS
	3		JAS
	4		AS
7	1	1/87	S

1986 Panel

W	R	Int M.	Data for
2	4	8/86	J
3	1		JA
	2		JAS
	3		JAS
	4		AS
4	1	1/87	S

1987 Q_3

1985 Panel

W	R	Int M.	Data for
8	4	8/87 (omit)	J

1986 Panel

W	R	Int M.	Data for
5	4	8/87	J
6	1		JA
	2		JAS
	3		JAS
	4		AS
7	1	1/88	S

1987 Panel

W	R	Int M.	Data for
2	4	8/87	J
3	1		JA
	2		JAS
	3		JAS
	4		AS
4	1	1/88	S

1987 Q_3 2 Panels

1 Panel matched with 1986. ($\lambda_1 = .5000$)

No Panel matched with 1985. ($\lambda_2 = 0$)

Attachment III (cont.)

Fourth Quarter October, November, December

1984 Q_4

1984 Panel

W	R	Int M.	Data for
4	2	11/84	O
	3		ON
5	4	1/86	OND
	1		OND
	2		ND
	3	4/86	D

1985 Panel

W	R	Int M.	Data for
1	2	2/85	OND
	3		ND
	4	4/85	D

1985 Q_4

1984 Panel

W	R	Int M.	Data for
7	2	11/85	O
	3		ON
8	4	1/86	OND
	1		OND
	2		ND
	3	4/86	D

1985 Panel

W	R	Int M.	Data for
3	3	11/85	O
	4		ON
4	1	1/86	OND
	2		OND
	3		ND
4	4	4/86	D

1986 Panel

W	R	Int M.	Data for
1	2	2/86	OND
3			ND
4	4	4/86	D

1986 Q_4

1985 Panel

W	R	Int M.	Data for
6	3	11/86	O
	4		ON
7	1	1/87	ON
	2		OND
	3		ND
	4	4/87	D

1986 Panel

W	R	Int M.	Data for
3	3	11/86	O
	4		ON
4	1	1/87	OND
	2		OND
	3		ND
	4	4/87	D

1987 Panel

W	R	Int M.	Data for
1	2	2/87	OND
	3		ND
	4	4/87	D

1987 Q_4

1986 Panel

W	R	Int M.	Data for
6	3	11/87	O
	4		ON
7	1	1/88	OND
	2		OND
	3		ND
	4	4/88	D

1987 Panel

W	R	Int M.	Data for
3	3	11/87	O
	4		ON
4	1	1/88	OND
	2		OND
	3		ND
	4	4/88	D

1988 Panel

W	R	Int M.	Data for
1	2	2/88	OND
	3		ND
	4	4/88	D

1987 Q_4 2.5 Panels

1.5 Panels matched with 1986. (λ_1 = .6000)

½ Panel matched with 1985. (λ_2 = .2000)