EXPANSION OF CPS TO PROVIDE RELIABLE STATE ESTIMATES OF UNEMPLOYMENT

Cathryn Dippo, Bureau of the Census

I. Introduction

The Comprehensive Employment and Training Act of 1973 (CETA) provides for 37½ percent of the funds available under Title I to be distributed to prime sponsors within the individual States on the basis of the "relative number of unemployed persons within the State as compared to such numbers in all States." (1)

In the past, the Bureau of the Census has produced estimates of unemployment for some States using the Current Population Survey (CPS) by merely tabulating sample units by States using their national weights. Since CPS was designed solely to produce national estimates, the reliability of these estimates varies considerably among the States. At the request of the Bureau of Labor Statistics (BLS), the Census Bureau has designed supplementary samples to provide better State estimates at a minimum level of reliability for all States. The criterion for the reliability of the State estimates set by BLS requires the annual average number of unemployed persons be provided with a coefficient of variation of ten percent or less for each State assuming an unemployment rate of six percent.

After changing the way State estimates are produced from the national file by introducing ratio estimation at the State level, only 23 States meet the reliability requirement. Therefore, it has been necessary to supplement the national sample in the remaining 27 States and the District of Columbia.

The purposes of this paper are twofold: (1) to present a general overview of the CETA supplemental samples, and (2) to discuss the computer program written to select the CETA samples. The program covers all phases of sample design and selection including estimation of sample size, stratification, PSU selection, and variance estimation of the resulting sample. This is the first time the Census Bureau has used computerized stratification for CPS.

From several alternative designs for the supplementation, the "dependent" procedure was chosen. A preliminary investigation indicated that be using this procedure smaller sample sizes would be required in most States to achieve the desired reliability. For this procedure the selection of the supplementary first stage sampling units is dependent upon the national sample PSU's. To provide a better understanding of the supplement, a brief description of the CPS sample design precedes the discussion of the CETA design.

II. CPS Sample Design

After the 1970 Decennial Census, CPS was redesigned to adjust for mobility of the population and any substantial changes in the characteristics of the population. The U.S. was divided into 1,924 primary sampling units (PSU's), each of which is a geographic area usually consisting of either a Standard Metropolitan Statistical Area (SMSA) or one or more contiguous nonmetropolitan counties not necessarily within the same State.

The PSU's were then grouped into 156 self-representing (SR) strata and 220 strata with more than one PSU (nonself-representing or NSR). In general, all PSU's with 250,000 or more inhabitants in 1970 were made SR. The remainder were combined with other PSU's in the same region maintaining, whenever possible, the stratification derived for the earlier CPS design.

The 220 strata were then grouped into 110 homogenous pairs. One stratum was picked at random from each pair of strata, each with equal probability. One PSU was then chosen from each selected stratum. From each of the remaining 110 strata, two PSU's were selected independently. Since the choices were independent, it was possible for the sample PSU's to be the same or different. The sample PSU's were selected with probabilities proportionate to size using a controlled selection procedure to maximize overlap with the old design. These selections resulted in 305 NSR PSU's for a total of 461 PSU's comprising 923 counties and independent cities. The monthly CPS sample consists of about 58,000 designated or about 45,000 interviewed households in these 461 PSU's.

The most important aspect of the CPS design, as related to the CETA design, is the way in which the NSR PSU's are stratified. For CPS, PSU's are combined with others in the same region using the following characteristics: SMSA-non-SMSA, percent of population living in urban areas, proportion of population nonwhite, per capita retail sales, rate of population change from 1960 to 1970, percent of population in manufacturing, and principal industries.

As a result, many strata contain PSU's from more than one State. Therefore, although there is a sample control requiring some representation in every State and the District of Columbia, the number of sample PSU's in any one State can be, and is, as low as one (Alaska, Hawaii, Nevada).

III. CETA Sample Design

For each of the 27 States and the District of Columbia, a supplemental sample (referred to as the CETA sample) has been designed which attempts to maximize the use of the national CPS sample. That is, all national CPS sample PSU's are also CETA sample PSU's. However, the CPS sample PSU's represent only that portion of the national CPS strata within the State. All PSU's not represented by CPS sample PSU's are then regrouped into strata within the State. As a result, a different set of national CPS sample PSU's would have generated a different CETA sample.

As in the national CPS, a two-stage stratified design has been used in most States. The five exceptions are Delaware, District of Columbia, New Hampshire, Rhode Island, and Vermont where each PSU in the State is in sample with certainty. This is because all of Rhode Island and the District of Columbia are SR in national CPS, and a large variance contribution from sampling PSU's in Delaware, New Hampshire, and Vermont would outweigh the cost advantage of concentrating the

sample within a few PSU's. For each of the remaining 23 States, all NSR PSU's for the State not represented by a national CPS sample PSU in the State have been restratified. PSU's are defined exactly as in the national CPS except for Hawaii and those which cross State boundaries. New PSU's have been defined in Hawaii to restrict the number of islands in sample each month thereby reducing interview costs.

The CETA sample units are drawn from within the sample PSU's using the same sampling frame as the national sample and the same systematic sampling procedures. The CETA sample units are also divided into eight rotation groups which rotate in the same way as the national sample. Since the overall sampling fractions for each of the CETA samples differ from that of the national CPS sample, CETA sample PSU's either contain too few or too many sample units for a State selfweighting sample. Where possible, additional sample selections within a PSU are made to avoid overlap with CPS. In PSU's containing more sample than required, all of the sample units are used in the State estimates but with an appropriate weight adjustment.

IV. Computer Program

In the past stratification for the national CPS has been done subjectively by combining PSU's that appear to be similar in selected characteristics. Since the stratification was done by hand, only a limited number of characteristics could be used and the choice of the characteristics to be considered was intuitive. To increase the number of stratification variables and the amount of objectivity, a stepwise multiple regression program has been used for stratification. Since the stratification of NSR PSU's not represented by CPS sample PSU's is only part of one step in the procedure for selecting the CETA State samples, the regression program has been converted to a subroutine and included in a program which executes the entire procedure as outlined below.

A. A preliminary estimate of the total number of designated households h needed in sample per month is made for the State using an estimate of the variance of the resulting sample. The variance can be crudely approximated by adjusting the simple random sample variance by a design effect.

The total effective annual sample size needed in persons to achieve a ten-percent coefficient of variation assuming simple random sampling is

$$m = \frac{q}{p(.1)^2}$$

where $\hat{u}$ = .05 $\underline{1}/_x$ 1970 Census State civilian labor force

= number of unemployed in 1970 if unemployment rate were five percent,

P = 1970 Census State population,

$p = \frac{\hat{u}}{P}$, and

q = 1-p.

Converting m into an equivalent number of yearly designated housing units assuming a noninterview rate of 15 percent yields

$$h = \frac{115\ qHF}{\hat{u}}$$

where H = total State housing units, and
F = design effect.

After applying the fpc factor and converting to monthly designated housing units by assuming an efficiency factor of .2 for an annual average of unemployment (2), the number of housing units to be designated per month is

$$h = \frac{\hat{h}\ H}{5(\hat{h}+H)} .$$

For the first trial estimate, the overall design effect F is assumed to be 1.75.

B. The estimated sample size h and the total State housing units determine a sampling fraction for a State selfweighting sample. This sampling fraction is then used to define a stratum size consistent with the desired interviewer workload of 50 to 55 designated sample units.[2/] PSU's with population greater than this stratum size are defined as self-representing. Since national CPS sample PSU's represent the portion of the national strata within the State, all PSU's which are SR for the national CPS sample are SR for the CETA sample. In addition any national NSR sample PSU which is the only PSU in its national stratum from the State becomes SR for CETA.

C. The remaining PSU's are assembled into strata and one NSR sample PSU is selected per stratum with probability proportionate to size subject to the following constraints.

1. All PSU's in a national stratum represented by a CPS sample PSU from the State must remain in the same stratum. That is, a CPS sample PSU is in sample for CETA, but the stratum it represents is the portion of its national stratum within the State.

2. Strata formed from the remaining PSU's, i.e., those not represented by a national sample PSU, should be of sufficient size to yield approximately 40 to 70[2/] sample households.

The strata are formed using the resulting predictive equation obtained from a multiple stepwise regression subroutine. Various economic and social characteristics for the NSR PSU's are regressed on the unemployment rates of the PSU's for each of the States using 1970 data from both the Decennial Census and County Business Patterns. The PSU's are then sorted by their predicted unemployment rates. The population of the PSU's is then accumulated until the combination of PSU's has a population sufficient to yield one interviewer workload. The process is repeated until all the PSU's are assembled into strata.

D. The components of the sampling error of the resulting stratification are then estimated. The between PSU component of variance is computed as a population variance using 1970 and 1960 Census unemployment distributions adjusted to yield a five percent unemployment rate. A design effect of 1.4 over simple random sampling is used to approximate the within PSU variance component. Together these components are substituted for the more crude estimate of variance made with the overall design effect F in IVA.

$$\text{Total variance} = \underbrace{kfpqP}_{\text{within PSU}} + \underbrace{c^2\left[\sum_h \sum_i \frac{NSR\ P_h}{P_{hi}} X_{hi}^2 - NSR \sum_h X_h^2\right]}_{\text{between PSU}}$$

where $k=\frac{H}{5h}$

$f=1.4$

P_h=stratum population as set up by above procedures

P_{hi}=PSU population

X_h=stratum unemployment

X_{hi}=PSU unemployment

$c_{70}=\frac{\hat{u}}{\text{1970 Census State unemployment 16+}}$

$c_{60}=\frac{\hat{u}}{\text{1960 Census State unemployment 14+}}$

The coefficients of variation for the 1970 and 1960 unemployment estimates are then computed. The 1960 coefficient of variation is used as an indication of how reliable the sample might be after ten years, and is expected to be larger than the 1970 coefficient of variation. If the larger of the two coefficients of variation is close to, but not over, ten percent, the sample design is acceptable. Usually, however, the entire procedure must be repeated several times with a different sample size to achieve an acceptable design. It is this repetitive aspect of the sample design which originally led to the use of a computer to execute this procedure.

V. Some Results

One hundred seventy new sample PSU's including rotating or replacement PSU's, have been designated for the CETA sample. However, the number of new areas in sample per month varies from 161 to 164 due to the rotation of PSU's between States for national CPS and, therefore, between CPS and CETA sample PSU's. The total number of areas in sample per month for both the national and State samples is about 623. Approximately 12,000 additional households are to be designated for sample per month bringing the total to 70,000. Assuming recent noninterview rates, about 54,000 households will be interviewed per month.

The distribution of the supplementary sample is shown in table 1, column 3. Columns 1 and 2 show the number of designated households available from CPS and how many of these households would have been selected for a State sample. Although all the households will be used with appropriate weight adjustments, the differences in these two numbers indicates some of the inefficiencies of the CPS design for making State estimates.

Estimates of the within and between PSU components of variance, coefficients of variation, and design effects for the CETA samples are given in table 2 for 1960 and 1970 Census unemployment distributions adjusted to yield a five-percent unemployment rate. The ratios in columns 6 and 7 indicate the relative importance of the between PSU component of variance of 1960 and 1970 respectively. Since the PSU's are stratified using 1970 data, the between PSU component of variance for 1970 is expected to be lower than it will be at the end of the decade before the next redesign. Therefore, the between PSU component of variance was also computed using the 1960 unemployment distribution to get an idea of how much the sample might deteriorate in a ten-year period.

The between PSU variance for 1970 is lower than for 1960 for most States. In fact, the between PSU variance for 1960 for Arizona, Utah, and Wyoming is as much as seven to ten times that for 1970. Of the three States for which the opposite is true - Arkansas, Idaho, and Iowa - the greatest difference is for Idaho where the between PSU variance for 1970 is almost double that for 1960.

The overall design effects of the resulting CETA samples, assuming a design effect of 1.4 for within sampling, are in columns 10 and 11 for 1960 and 1970 respectively. These design effects for the States range from 1.49 to 2.68 for 1960 data and from 1.42 to 2.16 for 1970 data. The average design effects are 1.92 for 21 States for 1960 and 1.65 for 1970. As a comparison, the design effect of CPS for the unbiased estimate of unemployment at the national level for 1965-1969 ranged from 1.43 to 1.56. (2)

As pointed out before, one major difference between the methods of designing the CETA samples and national CPS is in the method of stratification. For CPS, a card has been made for each PSU containing information on the characteristics thought to be important for stratification. The cards are then grouped together into strata of approximately the same population size trying to achieve homogeneity for these characteristics within strata, while maintaining the 1960 strata as much as possible. For CETA, various characteristics are regressed on unemployment to determine which had the highest correlations with unemployment and the interrelationship between the characteristics. PSU's are then grouped together using their predicted unemployment rates.

Although the specific set of characteristics used in the regression varies from State to State, the following are used for most States: the rate of population change for 1960 to 1970; the percent nonwhite; the percent males 18+ of total population; the ratio of workers to nonworkers; the proportion of population employed as white collar, blue collar, farm, and service workers; the percent of population living in urban, rural farm, and rural nonfarm areas; and the proportion of total employed in construction, manufacturing, transportation and public utilities, wholesale trade, retail trade, finance, services and professions, and medical services. Data for the principle industry variables are from the 1970 County Business Patterns. The remaining data are from the 1970 Decennial Census.

The first six variables entering into the regression equations for each State are given in order of entrance with the resulting multiple correlation coefficients (R) in table 3. There are a few States for which fewer than six variables entered the predictive equation; but for most States more than six entered. However, the reduction in residual variance is usually very small for the additional variables and, as expected, there are a few States for which the entire set of variables did not reduce the residual variance to

TABLE 1. Monthly Sample Sizes for CPS and CETA

(Designated Households)

STATE	Monthly CPS Sample		Additional Monthly Sample For CETA	Total Monthly Sample Size Needed For CETA (2) + (3)	Total Available Monthly Sample CPS + CETA (1)+(3)
	Total	Used in CETA[1/]			
	(1)	(2)	(3)	(4)	(5)
Alaska	90	90	920	1,010	1,010
Arizona	520	405	190	595	710
Arkansas	570	445	325	770	895
Colorado	520	420	445	865	965
Delaware	135	135	380	515	515
District of Columbia	215	215	285	500	500
Hawaii	135	135	390	525	525
Idaho	210	155	565	720	775
Iowa	615	330	345	675	960
Kansas	580	430	270	700	850
Maine	410	305	415	720	825
Minnesota	1,040	580	170	750	1,210
Mississippi	445	370	305	675	750
Montana	200	130	690	820	890
Nebraska	380	335	385	720	765
Nevada	70	70	565	635	635
New Hampshire	285	220	355	575	640
New Mexico	250	180	575	755	825
North Dakota	190	175	730	905	920
Oklahoma	765	580	160	740	925
Oregon	600	435	225	660	825
Rhode Island	240	240	270	510	510
South Carolina	570	340	170	510	740
South Dakota	180	130	850	980	1,030
Utah	170	170	710	880	880
Vermont	135	90	495	585	630
West Virginia	425	345	345	690	770
Wyoming	120	120	565	685	685
Total	10,065	7,575	12,095	19,670	22,160

1/ CPS occasionally provides more sample within a sample PSU then is actually needed for a self-weighting State sample. All sample units will be used but with the appropriate weight adjustment.

TABLE 2. Estimated Variances and Design Effects of CETA Sample for 1960 and 1970 Census Unemployment (Adjusted to Yield a Five-Percent Unemployment Rate)

STATE	Total Variance (000)		Within PSU 1/ (000)	Between PSU (000)		Ratio of Variances (%) Between/Total (4)/(1)	(5)/(2)	Coefficient of Variation (%)		Design Effect	
	1960	1970		1960	1970	1960	1970	1960	1970	1960	1970
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Alaska	NA2/	159	141	NA	18	NA	11.32	NA	8.12	NA	1.59
Arizona	10,147	9,196	9,099	1,048	97	10.33	1.05	9.94	9.46	1.56	1.42
Arkansas	11,307	11,583	8,712	2,595	2,872	22.95	24.79	9.77	9.88	1.82	1.86
Colorado	17,859	12,090	10,863	6,996	1,227	39.17	10.15	9.80	8.07	2.30	1.56
Idaho	1,580	1,835	1,329	251	506	15.89	27.57	9.26	9.98	1.66	1.93
Iowa	31,081	31,565	23,302	7,780	8,263	25.03	26.18	9.89	9.97	1.87	1.90
Kansas	19,206	16,102	14,425	4,781	1,677	24.89	10.41	9.89	9.05	1.86	1.56
Maine	3,619	3,130	3,032	588	99	16.25	3.16	9.97	9.27	1.67	1.45
Minnesota	58,113	50,075	37,414	20,699	12,662	35.62	25.29	9.98	9.26	2.17	1.87
Mississippi	14,250	14,011	11,338	2,912	2,672	20.44	19.07	9.98	9.90	1.76	1.73
Montana	1,689	1,617	1,132	577	485	32.98	29.99	9.97	9.76	2.09	2.00
Nebraska	8,689	7,557	6,103	2,587	1,455	29.77	19.25	9.96	9.29	1.99	1.73
Nevada	1,089	851	814	275	37	25.25	4.35	9.99	8.83	1.87	1.46
New Mexico	2,733	2,241	2,130	603	111	22.06	4.95	9.65	8.75	1.80	1.47
No. Dakota	1,119	859	699	420	161	37.53	18.74	9.87	8.65	2.24	1.72
Oklahoma	23,362	19,814	17,711	5,651	2,103	24.19	10.61	9.98	9.19	1.85	1.57
Oregon	17,126	14,174	13,625	3,501	549	20.44	3.87	9.89	9.00	1.76	1.46
So. Carolina	24,361	23,826	22,857	1,504	969	6.17	4.07	9.95	9.84	1.49	1.46
So. Dakota	1,506	1,282	829	677	453	44.95	35.34	9.84	9.08	2.54	2.16
Utah	3,960	2,206	2,066	1,894	140	47.83	6.35	9.97	7.44	2.68	1.49
West Virginia	8,065	7,639	7,239	827	401	10.25	5.25	9.80	9.54	1.56	1.48
Wyoming	418	325	316	102	9	24.40	2.77	9.98	8.80	1.85	1.44

1/ Approximated by assuming simple random sampling with a five-percent unemployment rate and a design effect of 1.4.

2/ 1960 Census data not available by 1970 Census districts.

Table 3. First Six Regression Variables in Order of Entrance and the Resulting Multiple Correlation Coefficient (R)

State	Variable 1	R	Variable 2	R	Variable 3	R	Variable 4	R	Variable 5	R	Variable 6	R
Alaska	Wholesale trade	.59	Total Manufacturing	.66	Blue Collar	.70	White Collar Worker-Non-Worker Ratio	.78	Worker Non-Worker Ratio	.82		
Arizona	Nonwhite	.78	Urban	.84	18+ Males	.90	Agricultural Svs.	.93	Construction	.95	Poverty	.98
Arkansas	Blue Collar	.48	Construction	.56	Hotels	.62	POP Change	.65	Apparel	.68	Rural Farm	.71
Colorado	Farm Workers	.55	Service Wkrs.	.64	Finance	.69	White Collar	.71	White Collar	.73	Rural Non-farm	.73
Idaho	Wood	.81	Rural Non-farm	.88	Farm Workers	.91	Rural farm	.92	White Collar	.93	Finance	.93
Iowa	Retail Trade	.51	Poverty	.56	Farm Workers	.72	Service Workers	.74	Finance	.75	Wholesale Trade	.77
Kansas	Farm Workers	.73	Machinery	.76	Service Wkrs.	.78	Total Mfg.	.79	Blue Collar	.80	18+ Males	.80
Maine	18+ Males	.77	POP Change	.95	Education	.97						
Minnesota	Wood	.52	Worker-Non-Wkr. Ratio	.58	Machinery	.63	Farm Workers	.67	Rural Farm	.73	POP Change	.77
Mississippi	Food	.40	Nonwhite	.54	Construction	.63	Medicine	.65	Service	.69	Rural Farm	.72
Montana	Farm Workers	.65	White Collar	.77	Worker-Non-Wkr Ratio	.80	Poverty	.82	Blue Collar	.84	Service Workers	.85
Nebraska	White Collar	.58	Urban	.62	Finance	.64	Farm Wkrs	.67	Retail Trade	.68	Food	.70
Nevada	Farm Workers	.60	Blue Collar	.86	18+ Males	.90	Total Manufacturing	.97	Rural Farm	.98	Medicine	.98
New Mexico	Farm Workers	.47	Public Utilities	.54	Rural Farm	.62	Service Workers	.65	Finance	.69	Total Manufacturing	.72
North Dakota	Rural Farm	.56	White Collar	.68	Blue Collar	.73	Food	.75	Total Mfg.	.78	Hotels	.79
Oklahoma	Farm Workers	.67	Poverty	.77	Medicine	.79	Oil & Gas	.81	Wholesale Trade	.82	POP Change	.83
Oregon	18+ Males	.51	Construction	.67	Worker -Non-Worker Ratio	.73	Wholesale Trade	.76	Rural Nonfarm	.78	White Collar	.82
So. Carolina	Worker-Non-Worker Ratio	.58	Apparel	.70	Service	.73	Rural Nonfarm	.76	White Collar	.79		
So. Dakota	Nonwhite	.76	Service Workers	.79	Service	.82	White Collar	.82	Farm Workers	.83	Poverty	.83
Utah	Retail Trade	.45	Public Utilities	.57	Service	.72	Nonwhite	.78	Construction	.80	Poverty	.83
West Virginia	Poverty	.57	POP Change	.73	Finance	.80	Wood	.82	Service Workers	.86	White Collar	.89
Wyoming	Hotels	.47	Wholesale Trade	.63	Medicine	.90	Oil & Gas	.81	Worker-Nonworker Ratio	.90	Retail Trade	.93

Table 4. Comparison of Between PSU Component of Variance For Computer and Hand Stratifications

State	# of New NSR PSU's	Between PSU Variance (000) Computer		Between PSU Variance (000) Hand		Coefficient of Variation Computer		Coefficient of Variation Hand		Ratio of Between PSU Variances Hand/Computer	
		1960	1970	1960	1970	1960	1970	1960	1970	1960 (4)/(2)	1970 (5)/(3)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Arizona	2	456	65	529	154	9.94	9.46	10.00	9.53	1.16	2.37
Arkansas	5	1168	1603	599	2164	9.77	9.88	9.54	10.07	.51	1.35
Colorado	3	2566	848	2601	1294	9.80	8.07	9.83	8.27	1.01	1.53
Idaho	4+1[1/]	257	539	738	1468	9.26	9.98	10.50	12.08	2.87	2.72
Iowa	6	1227	2552	2527	5388	9.89	9.97	10.46	10.95	2.06	2.11
Kansas	5	1661	603	1701	31647	9.89	9.05	9.91	18.59	1.02	52.48
Minnesota	3	8076	4290	11275	9878	9.98	9.26	10.54	10.44	1.40	2.30
Mississippi	5	1605	1439	2241	3184	9.98	9.90	10.18	10.51	1.40	2.21
Montana	5	835	518	928	732	9.97	9.76	10.14	10.18	1.11	1.41
Nebraska	4	473	221	583	305	9.96	9.29	10.15	9.46	1.23	1.38
Nevada	2	137	42	129	140	9.99	8.83	9.91	9.26	.94	3.33
New Mexico	7	674	139	567	273	9.65	8.75	9.49	8.94	.84	1.96
North Dakota	9	526	88	563	334	9.87	8.65	9.99	10.03	1.07	3.80
Oklahoma	3	1996	865	2316	1168	9.98	9.19	10.10	9.29	1.16	1.35
Oregon	2	1877	191	1922	195	9.89	9.00	9.90	9.00	1.02	1.02
South Carolina	2	604	325	927	973	9.95	9.84	10.09	10.08	1.53	2.99
South Dakota	11	354	197	414	429	9.84	9.08	10.13	10.47	1.17	2.18
Utah	3	262	29	383	159	9.97	7.44	10.32	7.65	1.46	5.48
West Virginia	2	689	34	613	45	9.80	9.54	9.79	9.55	.89	1.32
Wyoming	6	103	8	179	83	9.98	8.80	10.83	9.85	1.74	10.38

[1/] One CPS stratum had to be split because of size constraints. Different splits changed the PSU's to be stratified.

within tolerance. The proportion of variation explained by the first six variables varies from 71 to 98 percent. In the set of first six variables entering the regression equation, the proportion of farm workers occurs most frequently followed by white collar workers. Other important variables are the proportion of service workers, the worker-nonworker ratio, and the relative number of families in poverty.

In order to determine the relative effectiveness of the computer stratification method to the hand stratification method used for CPS, we have compared variances produced from the computerized stratification to those of a hand stratification similar to the type used for national CPS but based on the first few variables entering the regression. Table 4 shows the between PSU components of variance and coefficients of variation for both methods by State. The between PSU variance comparison has been made with Census data without the affect of adjusting unemployment rates to five percent.

Excluding Maine (which has only one new NSR strata) and Alaska (where computer stratification has not been used), there are 20 States for which a comparison can be made. The 1960 between PSU variance is lower for the hand stratification for four States; however, there are no States for which both the 1960 and 1970 between PSU variance is lower for the hand method.

Although for most States the gains of computer stratification are not substantial, this method of stratification is never shown to be worse than the old. Certainly this method is easier, faster, more versatile, and more objective. Although the program has been written to perform a specific task, it can be easily adapted to fulfill other constraints and, therefore, can be used for other supplements to CPS. Based on the results of the comparison of the two methods, further research is planned to examine the possibility of extending the use of computer stratification to the next CPS redesign.

1/ A five percent unemployment rate has been used for calculations to insure achieving the required reliability for a six percent rate.

2/ Based on production data for CPS interviewers in areas with low population densities, an optimum workload was estimated to be about 55 designated sample units Bounds on workload size were set at 40 and 70 units.

REFERENCES

(1) "Comprehensive Employment and Training Act of 1973," Public Law 93-203, 93rd Congress, S. 1559, December 28, 1973.

(2) Banks, Martha and Shapiro, Gary: "Variances of the Current Population Survey, Including Within- and Between-PSU Components and the Effect of the Different Stages of Estimation," Proceedings of the Annual Meeting of the American Statistical Association, Social Statistics Section, Fort Collins, Colorado, August 1975.