

Evaluation of Variance Estimation Methods Based on an SPD Longitudinal File

Smanchai Sae-Ung, David W. Hall, and C. Dennis Sissel, U.S. Census Bureau
 Contact Author Smanchai.Sae.Ung@census.gov

Key Words: variance estimation, replicate weights, balanced repeated replication, successive difference replication, residual technique of linearization, complex survey

Abstract: The following three methods are used, at the U.S. Census Bureau for direct variance and covariance estimation: the Fay's balanced repeated replication method, the combined Fay's balanced repeated replication and successive difference replication method, and the Deville's residual technique of linearization. In this study, we evaluate the differences in the variance and covariance estimates produced by these three methods for key characteristics of the longitudinal universe of the Survey of Program Dynamics represented by its longest longitudinal file. Based on the results of our evaluation, we identify various characteristics of the differences of the variance and covariance estimates produced by these three methods.

1. Introduction¹

Currently, at the U.S. Census Bureau, the following three methods are used for direct variance and covariance estimation in various complex demographic and socioeconomic surveys: the Fay's balanced repeated replication (BRR) method (Fay (1989)), the Fay's BRR method combined with the successive difference replication (SDR) method (Fay and Train (1995)), and the Deville's residual technique of linearization (DRL) method (Deville (1999)). Hereinafter, these three methods are simply referred to as the BRR method, the BRR-SDR method, and the DRL method, respectively. For examples, the Survey of Program Dynamics (SPD) and the Survey of Income and Program Participation (SIPP) use the BRR method for their public use micro data files, the Current

Population Survey (CPS) uses the BRR-SDR method for its public use micro data files, and The DataFerrett system is in the process of implementing the DRL method for the SIPP, as described by Rottach and Hall (2003). The DataFerrett System is the U.S. Census Bureau's online data extraction tool which also allows interactive data analysis. Under the current DataFerrett system, neither the BRR method nor the BRR-SDR methods are suitable for online interactive computation because of the sizes of their replicate weight files, while the DRL method is suitable, as demonstrated by Rottach and Hall (2003).

In this paper, we evaluate estimates of the variance and covariance produced by these three methods for key characteristics of the SPD longitudinal universe represented by the SPD third longitudinal file (the last and longest longitudinal file, covering 1992 to 2002). We determine the differences in the variance and covariance estimates produced by these three methods, and identify various characteristics of their differences.

2. SPD, Sample Design, and Longitudinal Weighting.

The SPD longitudinal universe consists of the civilian non-institutional population in the United States in March 1993. The survey universe is represented by the original sample people from the 1992 and 1993 SIPP panels (samples). The goal of the SPD is to provide policy makers a survey to assess the effect of the 1996 Welfare Reform Bill on the socioeconomic well being of the population. The SPD provides micro data covering a pre-reform through post-reform period from 1992 to 2002. The longitudinal nature is critical to conduct a pre/post analysis of the welfare reform in the nation and identifying interrelationships linking changes (such as welfare dependency and spells, employment, and child living environments).

¹ Disclaimer: This report is released to inform interested parties of research and to encourage discussion. The views expressed on statistical issues are those of the authors and not necessarily those of the U.S. Census Bureau.

The SPD sample was derived by sub-sampling from the 1992 and 1993 SIPP panels. The 1992 and 1993 SIPP panels were derived from the 1980 Sample Redesign (for demographic surveys at the U.S. Census Bureau). Their sample sizes were virtually the same. The SPD sample was derived such that approximately half of the sample was selected from the 1992 SIPP panel and the other approximate half was selected from the 1993 SIPP panel. A full description of the derivation of the SPD sample is available in SPD Technical Documentation (2004).

The 1992 and 1993 SIPP panels were designed using a two-stage probability sampling process. In the process, a sample of primary sampling units (PSUs) was selected, followed by a systematic sampling of housing unit clusters (a.k.a. ultimate sampling units (USUs)) from each sample PSU. (A USU generally consisted of two housing units in the address sampling frame, and four housing units in the area and new construction sampling frames.) The PSUs were classified into two types based principally on their measure of size (MOS). A PSU with larger MOS was classified as a self-representing PSU (SR PSU), and a PSU with smaller MOS was classified as non-self-representing PSU (NSR PSU). An SR PSU was selected with certainty. The NSR PSUs in each of the Census Regions (Northeast, South, Midwest, and West) were grouped into strata based on their demographic and socioeconomic similarity. Two NSR PSUs were selected from each stratum with the selection probability proportional to MOS. For each of the 1992 and 1993 SIPP panels, 86 SR PSUs were selected (with certainty) and 198 NSR PSUs were selected from the 99 NSR PSU strata.

There are ten longitudinal weights on the SPD third longitudinal file. Like the SIPP, each weight was produced by a weighting process to effectively reduce the biases due to non-response and improve coverage. In this study, we evaluated the three variance estimation methods based on the 2001 traditional longitudinal panel weight which is the most versatile among the ten weights. This weight covers data collected from 1992 to 2001. A sample person is classified as longitudinally interviewed for this weight if he/she was self or proxy interviewed in, or her/his data was imputed for, every SIPP and SPD collection schedule (e.g., a wave). Like the other nine weights, it was derived using (a) a non-interview weight adjustment procedure to

reduce bias due to non-response, and (b) a second stage weight adjustment procedure by matching with demographic benchmark estimates (controls) to improve the coverage and to generally reduce variances of estimates. A full description of the weighting procedure is available in SPD Technical Documentation (2004).

3. BRR Method

Like other replication methods, in the BRR method, a set of replicate weights is generated to be used for total variance and covariance estimation (total variance = between-PSU variance and within-PSU variance). The information needed for the replicate weight generation is available on the 1992 and 1993 SIPP panel files, and includes an NSR PSU variance stratum code for each NSR PSU stratum, a half sample code (HC) for each USU in an NSR PSU, an NSR PSU variance stratum code for each USU in an SR PSU, and a half sample code (HC) for each USU in each SR PSU, as specified in Waite (1992). Since we have 99 NSR PSU variance strata, we theoretically need to generate at least 100 replicate weights. In this study, we conservatively generated 164 replicate weights. In the BRR replicate weight generation, we first calculated a replicate factor, RFb_{ir} for each USU as shown in equation 1 below.

$$RFb_{ir} = \begin{cases} 1 + 0.5h_{ir} & \text{if } HC = 1 \\ 1 - 0.5h_{ir} & \text{if } HC = 2 \end{cases} \quad (1)$$

where i denotes the i^{th} NSR PSU variance stratum code for $i = 1$ to 99; r denotes the r^{th} replicate for $r = 1$ to 164; and h_{ir} = the element ir of a 164×164 Hadamard matrix, = -1 or +1. For each replicate, we took the product of the initial weight and its replicate factor of each USU as its initial weight, and then performed the weighting in the same manner as the full sample weighting to obtain the replicate weight. A full technical description for the BRR replicate weight generation is available in Sae-Ung, Hall, and Sissel (2004).

A variance estimate $\text{Var}(Y_0)$ of a characteristic Y of the survey universe based on the BRR replicate weights can be expressed as shown in equation 2 below.

$$Var(Y_0) = \frac{4}{164} \sum_{r=1}^{164} (Y_0 - Y_r)^2 \quad (2)$$

where Y_0 is the estimate of the characteristic Y using the full sample weight, and Y_r is the estimate of the characteristic using the r^{th} replicate weight for $r = 1$ to 164. A covariance estimate $Cov(Y_0, Z_0)$ between two characteristics Y and Z of the survey universe based on the BRR replicate weights can be expressed as shown in equation 3 below.

$$Cov(Y_0, Z_0) = \frac{4}{164} \sum_{r=1}^{164} (Y_0 - Y_r)(Z_0 - Z_r) \quad (3)$$

4. BRR-SDR Method

Like the BRR method, we also generated 164 replicate weights for the BRR-SDR method. In the BRR-SDR method, for each USU in each NSR PSU, its replicate weight factor was calculated in the same manner as described in the BRR method. Since an SR PSU was selected with certainty, it only produced within-PSU variance. Thus, the assignment of each USU in each SR PSU to an NSR PSU variance strata in the BRR method for estimating total variance and covariance would theoretically introduce a fictitious amount of between-PSU variance to the total variance and covariance estimates due to the SR PSU. Thus, in the BRR-SDR method, like the CPS, we use the SDR method (Fay and Train (1995)) to calculate separately the replicate weight factors (associated with the within-PSU variance) for the SR PSUs. The information needed for calculating the replicate weight factor for each USU in each SR PSU (per the SDR method) is available in the 1992 and 1993 SIPP panel files, and includes stratification PSU identification, master sampling interval PSU (also known as basic PSU component (BPC)), and sampling hit string number. Within each BPC, the closer the values of two hit string numbers, the more similar the demographic characteristics among the USUs in the two hit strings. Thus, for each SR PSU, we sorted its USUs in this manner, which resulted in a sequence of USUs, say {USU1 to USUn}, in which two adjacent USUs have the most similar demographic characteristics. For the calculation of the replicate factor, we derived an $n_{max}+1 \times 164$ rectangular matrix H^* from a 164×164 Hadamard matrix H , where n_{max} was the largest

number of USUs within an SR PSU. Since $n_{max} = 630$, we reused H several times to derive H^* in similar manner to the CPS, as specified in Sae-Ung, Hall, and Sissel (2004). We then calculated a replicate factor $RF_{s_{ir}}$, for USU i in each SR PSU per equation 4 below.

$$RF_{ij} = 1 + (2)^{\frac{-3}{2} h^*_{ij}} - (2)^{\frac{-3}{2} h^*_{(i+1)j}} \quad (4)$$

where i denotes USU i and row i of H^* for $i = 1$ to n , and r denotes replicate r and column r for $r = 1$ to 164, and h^*_{ir} denote the element ir of H^* , $= -1$ or $+1$.

Like the BRR method, for each replicate, we took the product of the initial weight and its replicate factor of each USU as its initial weight, and then performed the weighting in the same manner as the full sample weighting to obtain the replicate weight. A full technical description for the BRR replicate weight generation is available in Sae-Ung, Hall, and Sissel (2004).

The variance and covariance estimates based on the BRR-SDR method can be calculated using equations 2 and 3.

5. DRL Method

The DRL method is the residual technique as proposed by Deville (1999) and implemented by Rottach and Hall (2003). This is a linearization method similar to the familiar Taylor linearization technique found in many statistical texts, but one in which final (post-stratified) weights are used in lieu of sample design (base or initial) weights. The information needed for this method is available on the SIPP and SPD files, and includes public use stratum and half sample codes as well as various demographic variables used in the second stage weight adjustment procedure.

Consider a statistic \hat{T} , which uses a final weight w (for example a ratio or mean $\hat{T} = \sum wy / \sum wz$). The first step of the DRL is to transform this variable following a linearization approach but replacing design weights with final weights. In our ratio example, this would be the familiar $(y - \hat{T}z) / \sum wz$. Similar transformations for other forms of \hat{T} can be found in the statistical literature. This transformed variable is then regressed onto marginal categorical variables to

linearize the nonlinearity brought by using final weights instead of design weights.

The set of marginal categorical variables indicate to which population control a particular observation was raked during the second stage weight adjustment procedure (post-stratification). For SPD, we need marginals from a three-dimensional table since we raked to three sets of population controls. One dimension consists of sex by age categories of Hispanics, another dimension is categories of sex by race by household and family member type of the total population, and the last dimension is categories of sex by race by age of the total population. A full description of the raking process is available in the SPD Technical Documentation (2004).

The residuals from this weighted least squares regression of the transformed variables onto the marginal categorical variables become our true linearized variables (Stukel, Hidioglou, and Sarndal (1996)) from which variances, covariances, and degrees of freedom can be estimated. Let e_{hik} be the residual from the weighted least squared regression for the k-th observation in the i-th half sample within the h-th stratum. Then

$$\hat{V}(\hat{T}) = \sum_h \{\hat{E}_{h1} - \hat{E}_{h2}\}^2 \quad (5)$$

where $\hat{E}_{hi} = \sum_k w_{hik} e_{hik}$ is the weighted sum of the residuals in half sample i of stratum h. The approximate degrees of freedom for this variance estimate is given by

$$\hat{d}\{\hat{V}, \hat{T}\} = \frac{3\{\hat{V}(\hat{T})\}^2}{\sum_h \{\hat{E}_{h1} - \hat{E}_{h2}\}^4} - 2 \quad (6)$$

6. Evaluation

The demographic and socioeconomic characteristics related to low income and poverty are part of the key characteristics of the SPD longitudinal survey universe. In this study, we evaluated the three variance and covariance estimation methods based on some of these characteristics using the 2001 traditional longitudinal weight (described in section 2). We evaluated household-, family-, and person-level

characteristics. The statistics for our evaluation included those commonly used in the demographic and socioeconomic analyses, i.e., total, proportion, mean, percentile, and relative change estimators. Thus, our evaluation included both smooth and non-smooth estimators. We also based our evaluation on cohorts with moderate (several thousands) and small (several hundreds) sample sizes that are more sensitive to difference in methodologies used for variance and covariance estimation. For our evaluation, we expressed the variance and covariance estimates in terms of coefficient of variation (CV) and correlation coefficient (CC), respectively.

The results of our evaluation are tabulated in tables 1 to 4. Namely, they included the following characteristics for all races, blacks, Hispanics, and children under the age of 18 for 1993, 1994, 2000, and 2001: households receiving public assistance payments, female headed families, people in families with income below 150% of the poverty threshold, family size of female headed families, incomes of families of four, and relative changes and correlation of median income for families of four, and proportions of people in families with income below 150% of the poverty threshold, between two years.

For the smooth estimators like totals, proportions, and means (tables 1-3), the differences between CVs produced by the BRR and BRR-SDR methods mostly varied from -10% to +12% (occasionally exceeded this bound but did not exceeded $\pm 18\%$) and scattered about the zero difference line without any patterns as shown in Figure 1. The differences produced by the BRR and DRL methods varied from -12% to +9%, and also scattered about the zero difference line without any patterns as shown in Figure 2 which is similar to Figure 1. For the non-smooth estimators like the median and first quartile (table 3) and the relative change of median income (table 4), the differences between CVs produced by the BRR and BRR-SDR methods varied from -25% to +14%, and scattered about the zero difference line without any patterns similar to Figure 1. The differences between CVs produced by the BRR and DRL methods varied from -22% to +25%, and scattered about the zero difference line without any patterns similar to Figure 1. For median family income and the proportion of people living in low income families (table 4), the differences between CCs

produced by the BRR and BRR-SDR methods varied from -14% to 15% , and the differences between the BRR and DRL methods varied from $+4\%$ to $+10\%$, for CC levels of 0.6 or higher. For CC level of 0.4 or lower, the differences between CCs produced by the BRR and BRR-SDR methods mostly varied from -27% to 49% , with one case having a difference of 241% (a small sample size of several hundred Hispanics), and the differences between the BRR and DRL methods mostly varied from -6% to $+46\%$, with one case having a difference of -81% (a small sample size of several hundred).

Based on the above comparison, the differences in the standard error estimates produced by the BRR, BRR-SDR, and DRL methods were generally moderate ($\pm 12\%$) for smooth estimators, but the differences were larger for non-smooth estimators ($\pm 25\%$). For the correlation coefficient estimates, the differences were rather moderate ($\pm 15\%$) when the correlation coefficient levels were high (0.6 and above), but when the correlation coefficient levels were low (0.4 or below), the differences were high (-27% to 49%). The differences were occasionally extreme when the sample sizes were small (several hundreds) and the ethnicity is Hispanic. At present, we cannot theoretically or analytically determine which one of the three methods produces the most accurate variance and covariance estimates. Even though the BRR-SDR method calculates the variance and covariance contributed from the SR PSUs and NSR PSUs separately, we cannot analytically show that the separation of the SR PSUs and NSR PSUs in the BRR-SDR method will necessarily produce more accurate estimates than those of the other two methods.

7. Discussion and Conclusion

We evaluated the variance and covariance estimates produced by the BRR, BRR-SDR, and the DRL methods, using the SPD third longitudinal file (the longest and the last among the SPD longitudinal files). The BRR method is currently used in the SPD and the SIPP, the BRR-SDR method is currently used by the CPS, and the DRL method will be used on the DataFerrett for SIPP and possibly other surveys (suitable for online interactive computation) for direct variance and covariance estimation. Our evaluation is based on the following key characteristics of the SPD longitudinal universe: households receiving public assistance

payments, female headed families, people in families with income below 150% of the poverty threshold, family size of female headed families, incomes of families of four, and relative changes and correlation of median income for families of four, and proportion of people in families with income below 150% of the poverty threshold, between two years. Our evaluation of these characteristics was for all races, blacks, Hispanics, and children under the age of 18, for 1993, 1994, 2000, and 2001. We also based our evaluation on cohorts with moderate (several thousand) and small (several hundred) sample sizes that are more sensitive to difference in methodologies used for variance and covariance estimation. We evaluated commonly used smooth and non-smooth estimators, and we also evaluated the correlation coefficient estimates. Our smooth estimators included totals, proportions, and means, and our non-smooth ones included percentiles and relative changes.

We found that, for the smooth estimators, the differences in the standard error estimates produced by the three methods were generally moderate, i.e., within $\pm 12\%$. For the non-smooth estimators, the differences were generally not quite so moderate, i.e., $\pm 25\%$. For the correlation coefficient estimates, the differences were rather moderate ($\pm 15\%$) for correlation coefficient levels of 0.6 and above, but the differences were high (-27% to 49%) for correlation coefficient levels of 0.4 or below. For small sample sizes (several hundred), the differences occasionally were more extreme for the non-smooth estimators and correlation coefficient estimates associated with Hispanic cohorts.

At present, we cannot analytically show which one of the three methods is capable of consistently producing more accurate variance and covariance estimates than the other two methods. This is due to (a) the complexity of each of the three methods per se as dictated by the complex survey sample design and the weighting process, and (b) the variability (variance) of the variance and covariance estimators produced by each of the three methods per se.

Based on the above discussion, we conclude that we cannot determine which method among the three methods is preferred for the direct variance and covariance estimation for a complex demographic survey like the SPD or the SIPP. However, we must mention that the variance and

covariance estimates produced by one method may be substantially different from those produced by the other two methods such that the differences in the variance and covariance estimates are large enough to alter the conclusion of a statistical inference like a hypothesis testing.

References

Fay, R.E. (1989). *Theory and Application of Replicate Weighting for Variance Calculation*. Proceedings of the Survey Research Methods Section, American Statistical Association, 212-217.

Waite, P.J. (1992). *SIPP 92: Half-Sample Codes for Public Use Files*. Memorandum for Courtland, S. Demographic Statistical Division, U.S. Census Bureau.

Fay, R.E., and Train, G. F. (1995). *Aspects of Survey and Model-Based Postcensal Estimation of Income and Poverty Characteristics for States and counties*. Proceedings of Government Statistics Section, American Statistical Association, 154-159.

Stukel, D.M., Hidiroglou, M.A., and Sarndal, C-E. (1996). Variance Estimation for Calibration Estimators: A Comparison of Jackknife Versus Taylor Linearization. *Survey Methodology*, Vol. 22, No. 2, 117-125.

Deville, J. (1999). *Variance Estimation for Complex Statistics and Estimators: Linearization and the Residual Techniques*. *Survey Methodology*, Vol. 25, No. 2, 193-203.

Rottach, R.A., and Hall, D.W. (2003). Methods of Statistical Inference for the Survey of Income and Program Participation that are Suitable for an Online Data Analysis Tool. Proceedings of Survey Research Methods Section, American Statistical Association, 3561-3567.

SPD Technical Documentation (2004). Third Longitudinal File, Survey of Program Dynamics (SPD), Demographic Surveys Division, U.S. Census Bureau.

Sae-Ung, S., Hall, D.W., and Sissel, C.D. (2004). Procedure for Replicate Weight Generation for the SPD Longitudinal Files. Memorandum for Documentation. Demographic Statistical Methods Division, U.S. Census Bureau.

Acknowledgement

We would like to gratefully thank our branch chief Tracy Mattingly for her support in the course of this research.

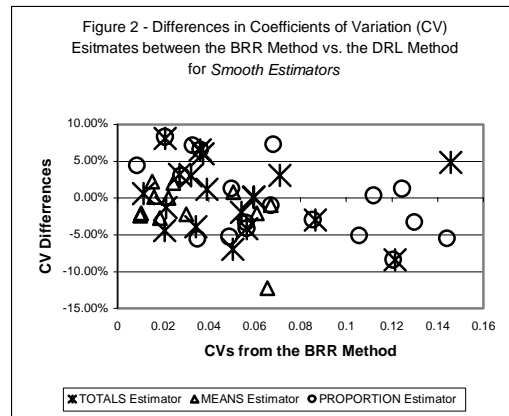
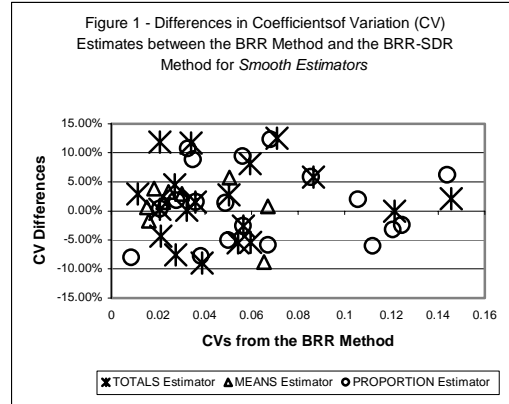


Table 1 – Evaluation of total estimators

Households receiving public assistance payments				
Year	Estimate	Method	CV	Difference from BRR*
1993	5,805,000	BRR	0.0361	
		BRR-SDR	0.0367	1.52%
		DRL	0.0386	6.58%
2000	2,007,000	BRR	0.0865	
		BRR-SDR	0.0902	5.82%
		DRL	0.0840	-2.94%
Black households receiving public assistant payments				
Year	Estimate	Method	CV	Difference from BRR*
1993	1,919,000	BRR	0.0565	
		BRR-SDR	0.0550	-2.59%
		DRL	0.0541	-4.15%
2000	716,000	BRR	0.1213	
		BRR-SDR	0.1213	0.00%
		DRL	0.1109	-8.38%
Hispanic households receiving public assistant payments				
Year	Estimate	Method	CV	Difference from BRR*
1993	1,267,000	BRR	0.0710	
		BRR-SDR	0.0798	12.47%
		DRL	0.0731	3.07%
2000	360,000	BRR	0.1456	
		BRR-SDR	0.1486	2.07%
		DRL	0.1384	4.89%
Female headed families				
Year	Estimate	Method	CV	Difference from BRR*
1993	14,872,000	BRR	0.0113	
		BRR-SDR	0.0116	2.97%
		DRL	0.0113	0.59%
2000	18,594,000	BRR	0.0212	
		BRR-SDR	0.0203	-4.36%
		DRL	0.0209	-1.24%
Black female headed families				
Year	Estimate	Method	CV	Difference from BRR*
1993	4,593,000	BRR	0.0207	
		BRR-SDR	0.0232	11.81%
		DRL	0.0198	-4.38%
2000	5,058,000	BRR	0.0373	
		BRR-SDR	0.0315	-15.37%
		DRL	0.0395	6.02%
Hispanic female headed families				
Year	Estimate	Method	CV	Difference from BRR*
1993	2,108,000	BRR	0.0594	
		BRR-SDR	0.0642	8.13%
		DRL	0.0599	0.075%
2000	2,581,000	BRR	0.0597	
		BRR-SDR	0.0565	-5.42%
		DRL	0.0598	0.17%
People in families with income below 150% of the poverty threshold				
Year	Estimate	Method	CV	Difference from BRR*
1993	47,113,000	BRR	0.0208	
		BRR-SDR	0.0209	0.24%
		DRL	0.0225	8.07%
2000	30,719,000	BRR	0.0271	
		BRR-SDR	0.0283	4.46%
		DRL	0.0280	3.23%
People in black families with income below 150% of the poverty threshold				
Year	Estimate	Method	CV	Difference from BRR*
1993	12,356,000	BRR	0.0342	
		BRR-SDR	0.0382	11.63%
		DRL	0.0329	-3.89%
2000	8,502,000	BRR	0.0505	
		BRR-SDR	0.0518	2.75%
		DRL	0.0470	-6.96%
People in Hispanic families with income below 150% of the poverty threshold				
Year	Estimate	Method	CV	Difference from BRR*
1993	11,605,000	BRR	0.0322	
		BRR-SDR	0.0359	0.13%
		DRL	0.0342	2.98%
2000	6,500,000	BRR	0.0543	
		BRR-SDR	0.0513	-5.53%
		DRL	0.0533	-1.91%
Children aged < 18 in families with income below 150% of the poverty threshold				
Year	Estimate	Method	CV	Difference from BRR*
1993	22,947,000	BRR	0.0275	
		BRR-SDR	0.0254	-7.58%
		DRL	0.0283	2.91%
2000	9,562,000	BRR	0.0390	
		BRR-SDR	0.0355	-9.00%
		DRL	0.0395	1.19%

* $((CV_{other} - CV_{BRR}) \div CV_{BRR}) \times 100$

Table 2 – Evaluation of proportion estimators

Households receiving public assistance payments				
Year	Estimate	Method	CV	Difference from BRR*
1993	6.03%	BRR	0.0362	
		BRR-SDR	0.0367	1.52%
		DRL	0.0385	6.58%
2000	1.57%	BRR	0.0855	
		BRR-SDR	0.0905	5.82%
		DRL	0.0830	-2.94%
Black households receiving public assistant payments				
Year	Estimate	Method	CV	Difference from BRR*
1993	17.08%	BRR	0.0565	
		BRR-SDR	0.0550	-2.59%
		DRL	0.0541	-4.15%
2000	4.84%	BRR	0.1206	
		BRR-SDR	0.1202	-3.20%
		DRL	0.1105	-8.38%
Hispanic households receiving public assistant payments				
Year	Estimate	Method	CV	Difference from BRR*
1993	16.07%	BRR	0.0681	
		BRR-SDR	0.0765	12.31%
		DRL	0.0732	7.32%
2000	3.36%	BRR	0.1440	
		BRR-SDR	0.1449	6.20%
		DRL	0.1361	-5.50%
Families of four with income below 150% of the poverty threshold				
Year	Estimate	Method	CV	Difference from BRR*
1993	18.24%	BRR	0.0500	
		BRR-SDR	0.0475	-5.11%
		DRL	0.0507	1.30%
2000	13.57%	BRR	0.0672	
		BRR-SDR	0.0632	-5.85%
		DRL	0.0665	-1.00%
Black families of four with income below 150% of the poverty threshold				
Year	Estimate	Method	CV	Difference from BRR*
1993	38.85%	BRR	0.1057	
		BRR-SDR	0.1078	1.97%
		DRL	0.1003	-5.10%
2000	26.73%	BRR	0.1245	
		BRR-SDR	0.1214	-2.43%
		DRL	0.1260	1.24%
Hispanic families of four with income below 150% of the poverty threshold				
Year	Estimate	Method	CV	Difference from BRR*
1993	43.10%	BRR	0.1120	
		BRR-SDR	0.1052	-6.07%
		DRL	0.1124	0.33%
2000	27.74%	BRR	0.1298	
		BRR-SDR	0.1121	-13.64%
		DRL	0.1256	-3.28%
People in families with income below 150% of the poverty threshold				
Year	Estimate	Method	CV	Difference from BRR*
1993	21.00%	BRR	0.0208	
		BRR-SDR	0.0208	0.34%
		DRL	0.0225	8.29%
2000	14.47%	BRR	0.0280	
		BRR-SDR	0.0285	1.80%
		DRL	0.0289	3.05%
People in black families with income below 150% of the poverty threshold				
Year	Estimate	Method	CV	Difference from BRR*
1993	43.01%	BRR	0.0349	
		BRR-SDR	0.0380	8.82%
		DRL	0.0330	-5.58%
2000	30.53%	BRR	0.0490	
		BRR-SDR	0.0497	1.35%
		DRL	0.0465	-5.24%
People in Hispanic families with income below 150% of the poverty threshold				
Year	Estimate	Method	CV	Difference from BRR*
1993	48.33%	BRR	0.0328	
		BRR-SDR	0.0363	10.72%
		DRL	0.0352	7.16%
2000	27.79%	BRR	0.0562	
		BRR-SDR	0.0509	-9.41%
		DRL	0.0543	-3.34%
Children aged < 18 in families with income below 150% of the poverty threshold				
Year	Estimate	Method	CV	Difference from BRR*
1993	31.17%	BRR	0.0085	
		BRR-SDR	0.0078	-8.07%
		DRL	0.0089	4.42%
2000	21.34%	BRR	0.0384	
		BRR-SDR	0.0354	-7.81%
		DRL	0.0396	3.05%

* $((CV_{other} - CV_{BRR}) \div CV_{BRR}) \times 100$

Table 3 – Evaluation of mean and percentile estimators

Mean size of female headed families				
Year	Estimate	Method	CV	Difference from BRR*
1993	2.829	BRR	0.0099	
		BRR-SDR	0.0115	17.55%
		DRL	0.0096	-2.41%
2001	2.391	BRR	0.0104	
		BRR-SDR	0.0118	13.35%
		DRL	0.0102	-2.06%
Mean size of black female headed families				
Year	Estimate	Method	CV	Difference from BRR*
1993	3.050	BRR	0.0222	
		BRR-SDR	0.0225	1.52%
		DRL	0.0207	6.45%
2001	2.636	BRR	0.0184	
		BRR-SDR	0.0191	3.85%
		DRL	0.0179	-2.71%
Mean size of Hispanic female headed families				
Year	Estimate	Method	CV	Difference from BRR*
1993	3.222	BRR	0.0301	
		BRR-SDR	0.0309	2.87%
		DRL	0.0294	-2.19%
2001	2.738	BRR	0.0245	
		BRR-SDR	0.0253	3.29%
		DRL	0.0250	2.06%
Mean income for families of four				
Year	Estimate	Method	CV	Difference from BRR*
1993	\$49,100	BRR	0.0152	
		BRR-SDR	0.0153	0.56%
		DRL	0.0156	2.27%
2000	\$70,100	BRR	0.0160	
		BRR-SDR	0.0157	-1.67%
		DRL	0.0160	0.19%
Mean income for black families of four				
Year	Estimate	Method	CV	Difference from BRR*
1993	\$36,500	BRR	0.0654	
		BRR-SDR	0.0596	-8.81%
		DRL	0.0573	-12.27%
2000	\$55,800	BRR	0.0506	
		BRR-SDR	0.0535	5.71%
		DRL	0.0510	0.80%
Mean income for Hispanic families of four				
Year	Estimate	Method	CV	Difference from BRR*
1993	\$31,700	BRR	0.0608	
		BRR-SDR	0.0504	-17.10%
		DRL	0.0595	-2.03%
2000	\$51,700	BRR	0.0670	
		BRR-SDR	0.0675	0.78%
		DRL	0.0663	-0.97%
Median income for families of four				
Year	Estimate	Method	CV	Difference from BRR*
1993	\$44,200	BRR	0.0220	
		BRR-SDR	0.0198	-10.27%
		DRL	0.0187	-15.00%
2000	\$63,700	BRR	0.0167	
		BRR-SDR	0.0190	14.19%
		DRL	0.0208	25.20%
Median income for black families of four				
Year	Estimate	Method	CV	Difference from BRR*
1993	\$29,000	BRR	0.1269	
		BRR-SDR	0.1277	0.61%
		DRL	0.1015	-20.05%
2000	51,400	BRR	0.0624	
		BRR-SDR	0.0619	-0.82%
		DRL	0.0782	25.22%
Median income for Hispanic families of four				
Year	Estimate	Method	CV	Difference from BRR*
1993	\$26,000	BRR	0.1602	
		BRR-SDR	0.1525	-4.80%
		DRL	0.1259	-21.42%
2000	\$43,800	BRR	0.0571	
		BRR-SDR	0.0523	-8.37%
		DRL	0.0618	8.21%
25 th percentile of income for families of four				
Year	Estimate	Method	CV	Difference from BRR*
1993	\$27,200	BRR	0.0266	
		BRR-SDR	0.0240	-9.68%
		DRL	0.0268	0.91%
2000	\$40,400	BRR	0.0259	
		BRR-SDR	0.0246	-5.23%
		DRL	0.0319	23.06%

* $((CV_{other} - CV_{BRR}) \div CV_{BRR}) \times 100$

Table 4 – Evaluation of relative change and correlation coefficient estimators (based on 1993 dollar value)

Relative change of median income for families of four				
Year	Estimate	Method	CV	Difference from BRR*
1993 and 1994	-0.27%	BRR	0.1118	
		BRR-SDR	0.0841	-24.78%
		DRL	**	
1993 and 2000	20.97%	BRR	0.0044	
		BRR-SDR	0.0033	-23.74%
		DRL	**	
Relative change of median income for black families of four				
Year	Estimate	Method	CV	Difference from BRR*
1993 and 1994	-6.71%	BRR	0.1328	
		BRR-SDR	0.1433	7.87%
		DRL	**	
1993 and 2000	48.98%	BRR	0.0739	
		BRR-SDR	0.0827	11.96%
		DRL	**	
Relative change of median income for Hispanic families of four				
Year	Estimate	Method	CV	Difference from BRR*
1993 and 1994	7.15%	BRR	0.2037	
		BRR-SDR	0.2172	6.65%
		DRL	**	
1993 and 2000	41.15%	BRR	0.1072	
		BRR-SDR	0.0880	-17.94%
		DRL	**	
Median income correlation coefficient for families of four				
Year	Method	Correlation Coefficient	Difference from BRR*	
1993 and 1994	BRR	0.5835		
	BRR-SDR	0.6112	4.74%	
	DRL	0.6347	8.78%	
1993 and 2000	BRR	0.1798		
	BRR-SDR	0.3615	101.10%	
	DRL	0.2145	19.50%	
Median income correlation coefficient for black families of four				
Year	Method	Correlation Coefficient	Difference from BRR*	
1993 and 1994	BRR	0.6760		
	BRR-SDR	0.6354	6.01%	
	DRL	0.6243	7.65%	
1993 and 2000	BRR	0.3789		
	BRR-SDR	0.2794	-26.26%	
	DRL	0.3555	-6.17%	
Median income correlation coefficient for Hispanic families of four				
Year	Method	Correlation Coefficient	Difference from BRR*	
1993 and 1994	BRR	0.6640		
	BRR-SDR	0.5707	-14.05%	
	DRL	0.7312	10.11%	
1993 and 2000	BRR	0.1874		
	BRR-SDR	0.2786	48.65%	
	DRL	0.0352	-81.22%	
Proportion of people in families with income below 150% poverty threshold				
Year	Method	Correlation Coefficient	Difference from BRR*	
1993 and 1994	BRR	0.6986		
	BRR-SDR	0.6493	-7.05%	
	DRL	0.7376	5.29%	
1993 and 2000	BRR	0.3030		
	BRR-SDR	0.3662	20.86%	
	DRL	0.4129	28.56%	
Proportion of people in black families with income below 150% poverty threshold				
Year	Method	Correlation Coefficient	Difference from BRR*	
1993 and 1994	BRR	0.7023		
	BRR-SDR	0.8097	15.29%	
	DRL	0.7300	3.79%	
1993 and 2000	BRR	0.4310		
	BRR-SDR	0.3155	-26.80%	
	DRL	0.4893	11.91%	
Proportion of people in Hisp. families with income below 150% poverty threshold				
Year	Method	Correlation Coefficient	Difference from BRR*	
1993 and 1994	BRR	0.6964		
	BRR-SDR	0.6822	-2.04%	
	DRL	0.7419	6.08%	
1993 and 2000	BRR	0.0644		
	BRR-SDR	0.2198	241.30%	
	DRL	0.1186	45.65%	

* $((CV_{other} - CV_{BRR}) \div CV_{BRR}) \times 100$

** This type of estimate currently not available in the DRL method software.