United States Department of Agriculture

National Agricultural Statistics Service

Research Division

SRB Research Report Number SRB-94-10

September 1995

REDRAWING THE 1994 FARM COSTS AND RETURNS SURVEY LIST FRAME SAMPLE TO REDUCE ITS OVERLAP WITH THE 1993 FCRS AND THREE OTHER MAJOR 1994 SURVEYS

Charles R. Perry
Jameson C. Burt
William C. Iwig

REDRAWING THE 1994 FARM COSTS AND RETURNS SURVEY LIST FRAME SAMPLE TO REDUCE ITS OVERLAP WITH THE 1993 FCRS AND THREE OTHER MAJOR 1994 SURVEYS, by Charles R. Perry, Jameson C. Burt, and William C. Iwig, Sampling and Estimation Research Section, Survey Research Branch, Research Division, National Agricultural Statistics Service, U.S. Department of Agriculture, Washington, D.C. 20250, September, 1995, Report No. SRB 94-10.

ABSTRACT

This report describes the Five Stage Algorithm used by the National Agricultural Statistics Service (NASS) to redraw the 1994 Farm Costs and Returns Survey (FCRS) list frame sample. This algorithm was devised to reduce the number of farm and ranch operators in the FCRS sample that are also in other major surveys' samples. It decreased the number of operators in the 1994 FCRS sample that were also in one or more of the four other major surveys from 4,338 to 1,982 (a decrease of 54 percent). It decreased the number of operators in the 1994 FCRS sample that were also in the previous year's FCRS sample from 369 to 14 (a decrease of 96 percent). It increased the number of operators in the 1994 FCRS sample that were not in any of the four other major surveys from 6,114 to 8,470 (an increase of 39 percent). We recommend caution when using this procedure, for example, using it on Ag surveys across years, or using presampling.

KEYWORDS

Respondent Burden, Multiple Selection, Cross-classification, Sample Coordination.

ACKNOWLEDGEMENTS

We would like to thank Ron Bosecker, George Hanuschak and Jim Davies for their continued support of our research into sampling methods that reduce burden. We would also like to thank Matt Fetter, and Wayne Dionne for their assistance in retrieving the large amount of data required for this application.

Table of Contents

SUMMARY

This report describes a sampling algorithm implemented by the National Agricultural Statistics Service (NASS) to reduce substantially the number of farmers selected for the 1994 Farm Costs and Returns Survey (FCRS) who were also selected for one or more of NASS's three other major 1994 surveys or for the 1993 FCRS. To accomplish this, NASS employed a five stage process.

At the first stage, the 1994 FCRS list frame sample was redrawn. The National Agricultural Statistics Service (NASS) used a two stage algorithm to redraw the 1993 Farm Costs and Returns Survey (FCRS) list frame sample as documented by Perry, Burt, and Iwig (1994). The algorithm was devised to reduce the number of farm and ranch operators in the 1993 FCRS sample that were also sampled for other major 1993 surveys and the 1992 FCRS, and thereby reduce respondent burden. The number of FCRS sampled operators included in other survey samples was substantially reduced from the initial FCRS sample to the final redrawn sample. A five stage algorithm was recommended by Perry, et al for redrawing the 1994 FCRS to reduce the potential for bias in the redrawn sample.

The report begins with a description of the five stage algorithm. It proceeds with theoretical and empirical justification for the algorithm, including tables showing that the estimated totals for eight control variables are very similar for the original and redrawn samples. Next, results are given. Tables are presented showing counts for different sample configurations of:

1) Records initially selected for the 1994 FCRS,

2) Records selected at the end of the first stage of the algorithm, and

3) Records selected for the final redrawn sample.

Following the results, conclusions are given. Among other encouraging findings, redrawing the 1994 FCRS sample:

1) Decreased the number of 1994 FCRS sample units that were in one or more of the other four surveys from 4,338 to 1,982 (a decrease of 54 percent),

2) Decreased the number of 1994 FCRS sample units that were also in the previous year's FCRS sample from 369 to 14 (a decrease of 96 percent), and

3) Increased the number of 1994 FCRS sample units that were not in any of the other four surveys from 6,114 to 8,470 (an increase of 39 percent).

The potential for bias resulting from the second stage of the algorithm in 1994 is discussed. It is concluded that any such bias will be much less than one percent of the 1994 FCRS estimates, hence undetectable in light of the coefficients of variation associated with the estimates.

We recommend redrawing the 1995 survey sample with the same 5 stage algorithm. Tables 4, 10, and 9 should be examined and estimates watched to insure that no biases occur.

Caution must be exercised when using these procedures. These approaches should not be used across years for the Ag Surveys, which are used to update the list frame. Presampling (prescreening) will not produce bias in the current year. However, it can produce bias in the next year unless those presampled are accomodated specially, for example, by treating them all as sampled in the previous year. A continuous classify would introduce bias.

Redrawing the 1994 Farm Costs and Returns Survey List Frame Sample to Reduce its Overlap with the 1993 FCRS and Three other Major 1994 Surveys

Charles R. Perry, Jameson C. Burt, William C. Iwig

INTRODUCTION

At the first stage, the 1994 FCRS list frame sample was redrawn. The National Agricultural Statistics Service (NASS) used a two stage algorithm to redraw the 1993 Farm Costs and Returns Survey (FCRS) list frame sample as documented by Perry, Burt, and Iwig (1994). The algorithm was devised to reduce the number of farm and ranch operators in the 1993 FCRS sample that were also sampled for other major 1993 surveys and the 1992 FCRS, and thereby reduce respondent burden. The number of FCRS sampled operators included in other survey samples was substantially reduced from the initial FCRS sample to the final redrawn sample. A five stage algorithm was recommended by Perry, et al for redrawing the 1994 FCRS to reduce the potential for bias in the redrawn sample.

This report describes how a special case of the Second-Method presented by Perry, Burt, and Iwig (1993) is used at each stage of the five stage process to redraw the sample. At the first stage the 1994 FCRS list frame sample is redrawn so that its overlap with the 1994 Quarterly Agricultural Survey (QAS), Agricultural Labor Survey (ALS), and Cattle and Sheep Survey (CSS) is minimized. In the next four stages, the redrawn sample is randomly shifted where necessary so that its overlap with the 1993 FCRS is minimized. The redrawing at each stage occurs among *population records* within a substratum. The substrata definitions change at each stage, but generally involve the intersection of survey strata, an age-of-control data variable, as a farm type variable. The result is a five stage algorithm to redraw the 1994 FCRS. That spreads the burden from multiple surveys as uniformly as possible over the population without changing the selection probabilities for the 1994 FCRS.

This report then describes the results of redrawing the 1994 FCRS sample. It gives the burden associated with the initial and with each of the succeeding five stages along with the burden reduction achieved. Also, tables are given that show, for eight population control variables, the estimated totals and coefficients of variation associated with the initial sample and the sample at each of the succeeding five stages.

This report next discusses any possible bias that the five-stage algorithm may introduce into the estimates. Particularly, it discusses biases that could result after the first stage of the algorithm. In addition, this report examines the potential for bias when applying this algorithm in two consecutive years.

Finally, this report makes its conclusions and recommendations for redrawing the 1995 FCRS.

In the remainder of this report, the qualifying phrase list frame is dropped from terms such as list frame sample. Since this report deals exclusively with the list frame part of NASS surveys, these abbreviations should cause no ambiguity.

FIVE STAGE ALGORITHM

Redrawing the 1994 FCRS used five stages, each using three basic steps. These three basic steps are described in the "Second-Method" of Perry, Burt, and Iwig (1993) for drawing samples for multiple surveys that minimize the burden on the individual units sampled:

Step 1. Use an equal probability of selection procedure within each stratum to select independent stratified samples for each survey.

Step 2. Cross-classify the population by the stratifications or other appropriate variables used in the individual surveys.

Step 3. Within each substratum, randomly reassign the samples associated with sampling units having excess burden to population units having less burden. Repeat the process until the burden on individual units sampled is minimized.

The five stage algorithm removes at each successive stage more and more of overlap between the 1993 and 1994 FCRS while maintaining at each stage, to the extent possible, the control-data's age and farm type characteristics of the sample. As a result, this should limit the potential for bias that might arise by applying the algorithm on two consecutive years (see Bias Considerations,).

Table 1 summarizes the variables which determine substrata at each stage. Sample records are only redrawn from other records within the same substratum. See Perry, Burt, and Iwig.

The first stage of the algorithm is the same as the first stage of the two-stage algorithm used last year. The substrata are based strictly on the intersection of the 1994 FCRS, QAS, ALS, and CSS strata. It removes, to the extent possible, the overlap of the 1994 FCRS sample with the other 1994 surveys' samples.

The second through fifth stages each use the Second-Stage of the Two-Stage Algorithm in Perry, Burt, and Iwig (1994b) with less and less restrictive age-dependent substrata. Two such age-dependent variables are formed, using an indicator or proxy for the age of control-data:

1. one indicator for crops which is the last update-year for any of Total Land In Farm, Total Cropland, or On-Farm Grain Storage control-values, and

2. one indicator for livstock which is the last update-year for any of All Cattle and Calves, Total Hogs and Pigs, or All Sheep control-values.

These two indicators are used in the second through the fourth stages, while their maximum value is used in the fifth stage. The second through fifth stages also restrict the substrata by FCRS Farm Type. This should help assure that, within the

Table 1. Variables Determining Substrata.

Variable	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
1993 FCRS stratum	—	yes	yes	—	—
1994 FCRS stratum	yes	yes	yes	yes	yes
1994 QAS stratum	yes	yes	—	—	—
1994 Labor stratum	yes	yes	—	—	—
1994 Cattle-Sheep stratum	yes	yes	—	—	—
1994 Farm Type	—	yes	yes	yes	yes
1994 age of crops	—	yes	yes	yes	—
1994 age of livestock	—	yes	yes	yes	—
1994 age of crops-livestock	—	—	—	—	yes

resulting substrata, the age indicators are more comparable.

The fourth and fifth stages do *not* restrict the substrata by the 1993 FCRS strata. The fifth stage is like the fourth stage, except it uses a single age indicator rather than two age indicators. Specifically, for an age indicator, the fifth stage uses the last year any of the following control-data was updated: Total Land In Farm, Total Crop Land, On-Farm Grain Storage, All Cattle and Calves, Total Hogs and Pigs, or All Sheep.

So, the second through fifth stages successively lower restrictions (and increase the size of the substrata) to reduce overlap between the 1994 and 1993 FCRS samples, at the cost of successively less protection against bias.

RESULTS

This section first gives the burden reduction at the U.S. and state level achieved by redrawing the sample. Included are Tables that show a detailed breakdown of the burden reduction achieved by redrawing the sample, along with the burden associated with the initial stage and the first through fifth stages.

Then this section discusses the similarity of the redrawn sample to the initial sample by examining the associated population control values. Tables are given that show for each of eight population control values the estimated totals and coefficients of variation that are associated with the initial and the first through fifth stages.

U.S. Level Burden Reduction

The reason for redrawing the 1994 FCRS sample was to reduce the burden on the

individual farmers selected by replacing the initial FCRS sample of 10,452 farmers with another sample of 10,452 farmers which places less burden on the individuals selected. In accomplishing this, the first stage of the algorithm reduced to the extent possible the overlap of the 1994 FCRS sample *with three other 1994 surveys' samples, the 1994 QAS, ALS, and CSS*. The second stage, while still using the 1993 FCRS strata to form substrata, formed the smallest substrata of any stage by crossing the stage one substrata with the 1993 FCRS and with the two age indicators. This second stage reduced to the extent possible the overlap of the 1994 FCRS sample *with the 1993 FCRS sample*, and did not seek to reduce overlap with the other three 1994 surveys. The later stages three through five formed larger substrata as indicated in Table 1. These stages three through five, like the second stage, only reduced to the extent possible overlap *with the 1993 FCRS sample*.

The burden reduction achieved by redrawing the 1994 FCRS can be summarized by the changes (or percentage changes) that occurred from the initial to the redrawn sample in the number of farmers selected for none, one, two, three, and four of the other surveys (1993 FCRS, 1994 QAS, ALS, and CSS).

From Table 2, under the second column "Sampling Configuration," the second configuration "10001" indicates farmers selected for both the 94 FCRS and the 94 CSS but for none of 93 FCRS, 94 QAS, or 94 ALS. In Stage 1, the algorithm sought to reduce such a configuration of surveys, reducing the number of farmers with this configuration from 889 initially to 537 at the end of Stage 1. Indeed, every row representing a combination of farmers in the 94 FCRS with any of 94 QAS, ALS or CSS has fewer farmers at the end of Stage 1 than initially. Only the two rows where the Stage 1 didn't try to reduce burden — first row, "10000" and the fifth row "11000" — had more farmers at the end of Stage 1.

Notice also from Table 2 that for each successive stage after Stage 1, every configuration of surveys having both the 94 FCRS and the 93 FCRS got the same or fewer farmers. At the same time, while Stage 1 reduced the number of farmers getting the 94 FCRS and any of 94 QAS, ALS or CSS, after Stage 1, the number of farmers who simultaneously got no 93 FCRS only increased in succeeding stages. For example, the number of farmers getting only the 94 FCRS and the 94 CSS, the configuration "10001" in the second row of Table 2, progresses as

889 537 555 582 586 588

After Stage 2, the algorithm only reduced the number of farmers receiving both the 1994 FCRS and the 1993 FCRS. So, for example, the number of farmers getting the fourth configuration "11000" decreases after Stage 1

158 243 120 28 10 5

This also means that the number of farmers getting the 94 FCRS, the 93 FCRS and at least one of 94 QAS, ALS or CSS decreases at every stage. For example, the number of farmers getting the 94 FCRS, the 93 FCRS and the 94 CSS, the configuration "11001", decreases at each stage as follows.

57 38 34 4 2 0

Summarizing Table 2 from the Initial Stage through the final Stage 5, redrawing the 1994 FCRS sample of 10,452 farmers:

1. Increased at each stage the number of 1994 FCRS samples that were not selected for any of the other four surveys from 6,114 to 8,470 (an increase of 39 percent),

2. Decreased at each stage the number of 1994 FCRS samples that were selected for only one of the other four surveys from 3,476 to 1,644 (a decrease of 53 percent),

3. Decreased at each stage the number of 1994 FCRS samples that were selected for exactly two of the other four surveys from 776 to 307 (a decrease of 61 percent),

4. Decreased at each stage the number of 1994 FCRS samples that were selected for exactly three of the other four surveys from 78 to 30 (a decrease of 62 percent),

5. Decreased at each stage the number of 1994 FCRS samples that were selected for all of the other four surveys from 8 to 1 (a decrease of 88 percent), and

6. Decreased at each stage the number of 1994 FCRS samples that were also in the previous year's FCRS sample from 386 to 14 (a decrease of 96 percent).

This burden reduction across stages is broken down for large farmers, strata 90 and above, in Table 5; and for small farmers, strata 89 and below, in Table 6.

For each state, Tables 3, 7 and 8 show, at each stage, the number of farmers who would receive both the 1993 FCRS and the 1994 FCRS. Notably, after Stage 5, four farmers in Arizona still received the FCRS in both 1993 and 1994. Also note from Table 3, that most of the redrawing of the 1994 FCRS samples off of the 1993 FCRS samples occurred after Stage 3, with 52 samples remaining as overlap between the two surveys. The remaining two stages reduced the overlap to 14 samples. This indidcates a three stage process could be used to simplify the process, reduce the potential for bias (as compared to the five stage process) and still provide considerable reduction in the operlap between the two FCRSs.

U.S. Level Control Variable Estimates

The similarity of the redrawn sample to the initial sample can be seen through Tables 4, 9 and 10. These tables show for each of eight population control values the estimated totals and coefficients of variation that are associated with the initial, first stage, and redrawn samples.

Table 4 gives the U.S. level estimates, which are derived from the 10,452 farmers in each of the three samples. Table 9 gives the U.S. level estimates for strata 90 and above, which are derived from the 3,588 strata 90 and above farmers in each of the three samples. And, Table 10 gives the U.S. level estimates for strata 89 and below, which are derived from the 6,864 strata 89 and below farmers in each of the three samples.

In each table, column 1 shows the names of control items. Column 2 shows the

Table 2. For All Strata, U.S.: The Number Of FCRS Samples and Percent of Total FCRS Samples by Sampling Configuration for the Initial, First Stage, Second Stage, Third Stage, Fourth Stage, Fifth Stage 1994 FCRS Samples.

Combination of Surveys (1)	Sampling Configuration† (2)	Count						Percent					
		Initial	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Initial	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
	9 9 **4** 3 9 9 9 **F** F 4 4 4 **C** C Q A C **R** R A L S **S** S S S S	(#) (3)	(#) (4)	(#) (5)	(#) (6)	(#) (7)	(#) (8)	(%) (9)	(%) (10)	(%) (11)	(%) (12)	(%) (13)	(%) (14)
94 FCRS alone	**1** 0 0 0 0	6114	8332	8392	8451	8465	8470	58.5	79.7	80.3	80.9	81.0	81.0
	Total	6114	8332	8392	8451	8465	8470	58.5	79.7	80.3	80.9	81.0	81.0
94 FCRS plus one	**1** 0 0 0 1	889	537	555	582	586	588	8.5	5.1	5.3	5.6	5.6	5.6
other survey	**1** 0 0 1 0	277	70	77	83	84	84	2.6	0.7	0.7	0.8	0.8	0.8
	1 0 1 0 0	2152	859	905	960	965	967	20.6	8.2	8.7	9.2	9.2	9.2
	1 1 0 0 0	158	243	120	28	10	5	1.5	2.3	1.1	0.3	0.1	0.1
	Total	3476	1709	1657	1653	1645	1644	33.3	16.4	15.8	15.8	15.7	15.7
94 FCRS plus two	**1** 0 0 1 1	58	25	25	28	28	28	0.6	0.2	0.2	0.3	0.3	0.3
other surveys	**1** 0 1 0 1	399	211	221	231	232	233	3.8	2.0	2.1	2.2	2.2	2.2
	1 0 1 1 0	142	32	34	41	43	43	1.4	0.3	0.3	0.4	0.4	0.4
	1 1 0 0 1	57	38	34	4	2	0	0.6	0.4	0.3	0.0	0.0	0
	1 1 0 1 0	22	5	5	2	0	0	0.2	0.1	0.1	0.0	0	0
	1 1 1 0 0	98	54	38	7	5	3	0.9	0.5	0.4	0.1	0.1	0.0
	Total	776	365	357	313	310	307	7.4	3.5	3.4	3.0	3.0	2.9
94 FCRS plus three	**1** 0 1 1 1	35	17	20	24	24	25	0.3	0.2	0.2	0.2	0.2	0.2
other surveys	**1** 1 0 1 1	3	2	2	1	1	1	0.0	0.0	0.0	0.0	0.0	0.0
	1 1 1 0 1	31	21	18	7	5	3	0.3	0.2	0.2	0.1	0.1	0.0
	1 1 1 1 0	9	3	3	1	1	1	0.1	0.0	0.0	0.0	0.0	0.0
	Total	78	43	43	33	31	30	0.8	0.4	0.4	0.3	0.3	0.3
94 FCRS plus four	**1** 1 1 1 1	8	3	3	2	1	1	0.1	0.0	0.0	0.0	0.0	0.0
other surveys	Total	8	3	3	2	1	1	0.1	0.0	0.0	0.0	0.0	0.0
Total		10,452	10,452	10,452	10,452	10,452	10,452	100.0	100.0	100.0	100.0	100.0	100.0

† Note: The **Sampling Configuration** indicates the sampling pattern for the 94FCRS, 93FCRS, 94QAS, 94ALS and 94CSS (Cattle and Sheep). For example, a 10001 indicates sample units selected for only the 94FCRS and 94CSS.

Table 3. For All Strata: The Number of 1994 FCRS Samples by State that were in the 1993 FCRS Sample at each Stage of the Redrawing Process.

State	Overlap					
	Initial	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
01: Alabama	10	8	3			
04: Arizona	27	24	20	9	6	4
05: Arkansas	2	5	1	1	1	
06: California	29	24	18	4	2	
08: Colorado	9	9	4	1	1	
09: Connecticut	2	1				
12: Florida	15	23	10			
13: Georgia	10	15	11	1		
16: Idaho	8	9	5	2	2	1
17: Illinois	28	30	16	3		
18: Indiana	6	6	5			
19: Iowa	20	19	8			
20: Kansas	8	5	2	1		
21: Kentucky	8	9	7			
22: Louisiana	9	9	5	3	3	2
23: Maine	6	3	1			
24: Maryland	4	3	2	1		
25: Massachussetts	1	1				
26: Michigan	11	11	6	2		
27: Minnesota	29	19	14	3		
28: Mississippi	9	14	8	2		
29: Missouri	2	2	1	1		
30: Montana	3	2				
31: Nebraska	7	4	2			
32: Nevada	2	1	1			
33: New Hampshire	3	3	2	2	2	2
34: New Jersey	2	1	1			
35: New Mexico	7	5	4	2	2	
36: New York	8	7	2			
37: North Carolina	9	10	5	1		
38: North Dakota	6	4				
39: Ohio	7	7	6	3		
40: Oklahoma	3	2	1			
41: Oregon	10	8	6	2	1	1
42: Pennsylvania	3	1	1			
44: Rhode Island	3	2	2	2	1	1
45: South Carolina	6	6	3	1	1	1
46: South Dakota	7	8	8			
47: Tennessee	6	6	4			
48: Texas	16	15	9	3	1	
49: Utah	3	3	3			
50: Vermont	4	1				
51: Virginia	3	5	4			
53: Washington	8	11	6	1	1	1
54: West Virginia	1					
55: Wisconsin	3	4	2			
56: Wyoming	3	4	4	1	1	1
Total	386	369	223	52	25	14
Cumulative Percent Reduction		4.4%	42.2%	86.5%	93.5%	96.4%

1994 population totals for the control items. Columns 3 and 4 show, respectively, the lower and upper bounds for a 90 percent confidence interval centered on the true total for the estimated totals. Columns 5 through 10 show the estimated totals using the initial, Stage 1, ... and Stage 5 samples. Column 11 shows the coefficients of variation for the estimated totals. Columns 12 through 17 show the estimated coefficients of variation using the initial, Stage 1, ... Stage 5 samples.

These tables show that the three samples are very similar with respect to the estimates derived from them. At every stage, all the estimated totals in columns 5 through 10 fall within the confidence intervals of columns 3 and 4, with the exception of "Value of Sales." Moreover, all the estimates from the initial through Stage 5 tend to fall at almost the same place in their respective confidence intervals. For the "Value of Sales," the initial sample estimates a total just above the lower confidence bound 156.24 while the later samples estimate a total just below this lower confidence bound. In particular, the initial sample estimated total "Value of Sales" at 156.36 million dollars while the final Stage 5 sample estimated 155.58 million dollars, only 0.5 percent lower than the initial estimate.

The same is true for the estimated coefficients of variation. If the estimated coefficient of variation for an item is small relative to the true coefficient of variation for one sample then it is also small for the other two samples. The estimated coefficient of variation for All Cattle and Calves, which was 2.49 for the initial sample, increased to 3.63 for all later stages 1 through 5. Since the first stage redraws to reduce the number of farmers getting the 94 FCRS and some of 94 QAS, ALS or CSS, this increase can be attributed to random sampling. The increase in the estimated CV for all Cattle and Calves from the initial sample to the first stage in the 1994 redraw was 3.05 to 3.21 (Perry, Burt, Iwig (1994b), Table B16).. While the exact coefficients of variation in column 11 can be large, all the estimated coefficients of variation are small and consistent across stages.

BIAS CONSIDERATIONS

While the five-stage resampling algorithm can bias estimates, the size of the bias should be relatively small. For a sample to become biased, it must violate the assumption: *In any survey's strata, all farmers must be sampled at the same rate.* If previous survey data are used to update the frame and the sampling procedure selects records that were in the previous surveys at a lower (or higher) rate than at random, then bias can exist. All farmers within each stratum are *not* sampled at random. The bias occurs if the differences in the age of the control data for the records that are redrawn in stages two through five are related to this survey data collected from the records.

As noted by Perry et all (1994b), the first stage of the five-stage algorithm cannot bias the FCRS estimates. In this stage the 1994 FCRS sample records are moved off of the 1994 WAS, ALS, and CSS sample records. However, this is not related in any way to whether the control data were updated from 1993 surveys.

The second through fifth stages can introduce bias into the FCRS estimates since

information from the 1993 FCRS, QAS, ALS, and/or CSS was very likely used to update the 1994 FCRS frame. Only 355 out of 10,452 farmers were resampled in stages 2 through 5, representing about three percent of the FCRS sample. Last year, such farmers represented about 5 percent of control data totals. In the second through fifth stages, 1994 FCRS records were moved off of 1993 FCRS records. Since a similar two stage procedure was used last year (Perry et al 1994b), the 1993 FCRS records were separated from the 1993 QAS, ALS, and CSS sample records as much as possible. Therefore by moving off of the 1993 FCRS, the 1994 FCRS sample tends to contain more 1993 QAS, ALS, and CSS records than expected at random. And these records have updated control data. So sample records were not selected strictly at random.

Theoretically, it would be simple to test the second and latter stages of these algorithms for potential bias by either 1) collecting data from the farmers in the FCRS year-to-year overlap in addition to the data from redrawn farmers or 2) by splitting the FCRS year-to-year overlap and only using half of it. To split the year-to-year FCRS overlap one would divide the records that were redrawn at the second and latter stages into pairs based on all the available information, then randomly choose one record from each pair to be redrawn. The other record from each pair that would be used would be the record selected at the end of the first stage of the algorithm. All tests would need to be based on expanded data, since our interest is in potential for bias with respect to the FCRS estimates. Thus, in practice, neither procedure would likely yield useful information, since at best one would have no more than 400 records, which would be scattered over all strata in all states.

The use of age indicators in stages 2 through 5 to form substrata inhibits any bias since, in general, farmers will be resampled only among farmers with control data updated in the same year. This should make the control data for individual items comparable. In stages 2 through 5, restricting the substrata to one farm type should help make sure that the age indicators are more comparable within the substrata. This should further limit the potential for bias that might arise because of application of the algorithm in two consecutive years. These stages 2 through 5 decreased the possibility of bias from a strictly two stage algorithm as used in last year's resampling of the 1993 FCRS.

To show that the effect of this potential bias on the 1994 FCRS estimates is small, consider a (highly unlikely) worst-case example. Suppose that the change in age of the control-data causes the survey data collected from the part of the sample that is redrawn in stages 2 through 5 to be consistently 20 percent smaller (or larger) than the data that would have been collected had the elements not been redrawn. In reality, it is unlikely that slightly older or newer control-data would produce such a large difference in the relationship between the control data and observations.

From the discussion above, the expanded control-data of the three percent of the sample that was redrawn in stages 2 through 5 was about 5 percent. Making a worst case assumption, assume that the expanded data of the part of the sample redrawn in stages 2 through 5 represents five percent of the estimate. Even under these extreme assumptions,

redrawing the sample in stages 2 through 5 would cause only a one percent change ($0.20 * 0.05 = 0.01$) in the 1994 estimates. This change would be undetectable in light of the coefficients of variation associated with the FCRS estimates. Thus, even though the second through fifth stages of the algorithm cannot be guaranteed to be unbiased, it is highly unlikely that any detectable bias will result from redrawing the 1993 FCRS sample.

Since the estimated totals and estimated coefficients of variation hardly changed from the initial sample to the final sample, the five stage algorithm introduces no meaningful bias.

When one violates the equal-sampling-rate assumption in other ways, such small biases can become large biases. In particular, in a single stratum, when farmers are directly sampled at different rates, then estimates can be weighted or alternatively estimates can concoct new strata formed by the differing sampling rates. Such adjustments would produce unbiased estimates in the current year. But in the following year, such adjustments would not be accommodated by the five stage algorithm. This follows because within a stratum, the algorithm approach would insidiously substitute more for those farmers once sampled at a high rate than for those once sampled at a low rate. In the worst case, the previous year would sample at two rates in some stratum—100 percent and $<$ 100 percent—but the following year would then sample none of the 100 percent farmers. While the current year's estimates would be unbiased, the following year's would be biased. Such biases from unequal sampling can be accomodated by either (1) excluding from resampling any previous year's stratum of farmers which sampled within that stratum at more than one sampling rate, or (2) if subsampling created the unequal sampling rates, mark each farmer sampled at the first level as sampled, even when not sampled at the second level.

APPLICATIONS WITH PRESCREENING

For the 1995 FCRS, sample units in strata designated for either the sorghum or burley tobacco Cost of Production Survey (COPS) questionnaire were prescreened to improve the efficiency of the sample. With prescreening, relatively large initial samples are selected in the COPS strata. A telephone survey is then conducted in the fall to screen these sample units. Units that have the commodity of interest are selected for the COPS in the second phase sample. Units that do *not* have the commodity of interest are subsampled for other FCRS questionnaires. So sample units in the original COPS strata are sampled at different rates in the second phase for the actual FCRS/COPS.

There are two different options for applying the five-stage algorithm to redraw the current year sample when prescreening is used. These are:

1. redraw the *initial* sample in all strata including the COPS strata, or

2. redraw the second phase sample after the screening, within the positive substratum and the zero substratum of each original COPS stratum (if all records are selected in the positive substratum for the second phase sample, then there will be no redrawing within that substratum).

The redrawing can occur before the prescreening (option 1) or after prescreening (option 2). Option 1 was used for the 1994 FCRS. Redrawing the second phase sample for the COPS strata (option 2) would probably provide little reduction in respondent burden bacause the redrawing would be limited to within the positive and zero substrata of the original sample. All remaining population units in each COPS stratum which are not sampled in the initial screening phase would be placed in a third substratum. This substratum would be excluded from the redrawing process because it contains no FCRS sample units and it is unknown for each record whether it has the item of interest or not.

In the future, if prescreening was used in the previous year (as will be the case for 1996) it is critically important that the previous year FCRS sample within the COPS strata be treated in one of two ways. The intent is to insure that all population records within a strata (or substrata) for the current survey are sampled with equal probability. Consequently, we need to either redraw the current sample to remove duplication with the initial previous year sample in the COPS strata, or with the final (second phase) sample within second phase substrata. The cells used for redrawing would then only contain records selected with equal probability for the prevous year survey. If the cells contained records selected with unequal probability for the previous year, the algorithm would substitute more for those records sampled at the higher rate last year than for those records sampled at the lower rate. This implies that cells should not be defined by the initial previous COPS strata if only records in the final second phase sample are identified. More specifically, the two ways are:

1. Redraw within cells defined by the original previous year COPS strata *and* treat the initial COPS strata samples from the previous year as if they were the final samples. That is, do *not* distinguish between records that were screened to have the commodity of interest and those that do not have the commodity.

OR

2. Redraw within cells defined by the second phase substrata from the previous year COPS strata. The cells used for redrawing would then only contain records for a specific COPS substratum and all records would be redrawn with equal probability relative to the previous year sample.

CONCLUSIONS

Redrawing the 1994 FCRS list frame sample of 10,452 farmers reduced the burden on the individual farmer sampled by:

1. Decreasing the number of 1994 FCRS samples that were in one or more of the other samples (1994 QAS, ALS, CSS or 1993 FCRS) from 4,338 to 1,982 (a decrease of 54 percent),

2. Decreasing the number of 1994 FCRS samples that were in the 1993 FCRS sample from 369 to 14 (a decrease of 96 percent), and

3. Increasing the number of 1994 FCRS samples that are not in any of the other four surveys from 6,114 to 8,470 (an increase of 39 percent).

The five-stage algorithm used to redraw the 1994 FCRS has very little potential to introduce bias into the 1994 FCRS estimates. Only records redrawn in stages 2 through 5, which involve three percent of the sample, have any potential for introducing bias. Even if the survey data for these records was consistently 20 percent smaller (or larger) than this data would have been had the elements not been redrawn, the effect on the estimates would be only about 1 percent.

Hence, in comparison to the coefficients of variation associated with FCRS items, any bias should be undetectable. When using the five stage algorithm, Tables 4, 9 and 10 should be watched. In particular, columns 12 and 17 should not differ a great deal.

RECOMMENDATIONS

Redraw the 1995 and later FCRS samples with the five stage algorithem. For this five stage algorithm, consider any farmer sampled in a FCRS or COPS screening sample as sampled for the FCRS or COPS. After redrawing the sample, examine Tables 4, 9 and 10, paying particular attention to columns 12 and 17 for any large changes and possible biases.

REFERENCES

Perry, Charles R., Jameson C. Burt, and William C. Iwig; "Methods Of Selecting Samples In Multiple Surveys To Reduce Respondent Burden"; *Proceedings of the International Conference on Establishment Surveys*; ASA; Alexandria, Virginia; June 1993.

Perry, Charles R., Jameson C. Burt, and William C. Iwig; "Redrawing the 1993 Farm Cost and Returns Survey List Sample to Reduce Overlap with Three Other 1993 Surveys and the 1992 FCRS"; *1994 Proceedings of the Section on Survey Research Methods*; ASA; Alexandria, Virginia; 1994.

Perry, Charles R., Jameson C. Burt, and William C. Iwig; "Redrawing the 1993 Farm Cost and Returns Survey List Frame Sample to Reduce Its Overlap with the 1993 FCRS and Three other Major 1993 Surveys"; U.S. Department of Agriculture; Washington, D.C.; 1994.

APPENDIX

Table 4. For All Strata, U.S.: The Estimated 1994 Population Control Varia
Initial, First Stage, Second Stage, Third Stage, Fourth Stage, and Fifth Stage 19

Control Variable†	Pop. Total	90% Lower Bound	90% Upper Bound	Estimated Total						Exact CV of Estimated Total	Esti		
				Initial Sample	Stage 1 Sample	Stage 2 Sample	Stage 3 Sample	Stage 4 Sample	Stage 5 Sample		Initial Sample	Stage 1 Sample	Stag Sam
(1)	(#) (2)	(#) (3)	(#) (4)	(#) (5)	(#) (6)	(#) (7)	(#) (8)	(#) (9)	(#) (10)	(%) (11)	(%) (12)	(%) (13)	(% (14
Total Land in Farm	1494.84	-2641.81	5631.49	802.35	842.09	839.68	845.77	845.09	845.20	167.71	5.21	6.27	6.
Value of Sales	158.55	156.24	160.86	156.36	155.77	155.60	155.56	155.60	155.58	0.88	0.77	0.80	0.
Total Cropland	358.12	288.60	427.65	348.81	347.90	347.13	347.29	347.31	347.26	11.77	1.20	1.26	1.
On-Farm Grain Storage	998.63	187.52	1809.75	963.00	959.64	960.79	955.56	956.20	955.92	49.23	5.01	5.42	5.
All Cattle & Calves	864.17	809.11	919.23	809.51	837.34	838.32	838.67	838.87	838.90	3.86	2.49	3.63	3.
Total Hogs & Pigs	591.01	467.10	714.93	562.60	552.29	551.18	551.61	551.81	551.50	12.71	5.80	5.87	5.
All Sheep	104.07	72.11	136.03	116.49	117.77	117.81	117.79	117.78	117.89	18.61	15.87	15.74	15.
Farm Workers Hired	142.24	51.48	233.00	123.02	120.56	120.57	119.39	119.40	119.39	38.67	8.00	7.91	7.

† All missing control values were set to zero before any computations were performed. The units of measure
and their estimates are:

1,000,000 acres for Total Land;
1,000,000 acres for Total Cropland;
100,000 head for Cattle & Calves, Hogs & Pigs, and Sheep;
$1,000,000,000 for Farm Value Of Sales;
10,000,000 bushels for On-Farm Grain Stor
10,000 workers Farm Workers Hired.

Table 5. For Stratum 90 and Above, U.S.: The Number of FCRS Samples and Percent of Total FCI
Configuration for the Initial, First Stage, Second Stage, Third Stage, Fourth Stage, Fifth Stage 1994 FCR

Combination of Surveys	Sampling Configuration†	Count						Percent					
		Initial	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Initial	Stage 1	Stage 2	Stage 3	Stage 4	Stag
	9 9 **4** 3 9 9 9 **F** F 4 4 4 **C** C Q A C **R** R A L S **S** S S S S												
(1)	(2)	(#) (3)	(#) (4)	(#) (5)	(#) (6)	(#) (7)	(#) (8)	(%) (9)	(%) (10)	(%) (11)	(%) (12)	(%) (13)	(% (14
94 FCRS alone	**1** 0 0 0 0	1315	2091	2122	2160	2170	2174	36.6	58.3	59.1	60.2	60.5	60.
	Total	1315	2091	2122	2160	2170	2174	36.6	58.3	59.1	60.2	60.5	60.
	1 0 0 0 1	425	360	370	402	407	408	11.8	10.0	10.3	11.2	11.3	11.

Table 6. For all Stratum 89 and Below, U.S.: The Number of FCRS Samples and Percent of Total FC
Configuration for the Initial, First Stage, Second Stage, Third Stage, Fourth Stage, Fifth Stage 1994 FCF

Combination of Surveys (1)	Sampling Configuration† (2)	Count						Percent					
		Initial	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Initial	Stage 1	Stage 2	Stage 3	Stage 4	Stag
	9 9 4 3 9 9 9 F F 4 4 4 C C Q A C R R A L S S S S S S	(#) (3)	(#) (4)	(#) (5)	(#) (6)	(#) (7)	(#) (8)	(%) (9)	(%) (10)	(%) (11)	(%) (12)	(%) (13)	(% (14
94 FCRS alone	**1** 0 0 0 0	4799	6241	6270	6291	6295	6296	69.9	90.9	91.3	91.7	91.7	91.
	Total	4799	6241	6270	6291	6295	6296	69.9	90.9	91.3	91.7	91.7	91.
94 FCRS plus one other survey	**1** 0 0 0 1	484	171	176	179	179	180	7.0	2.5	2.6	2.6	2.6	2.
	1 0 0 1 0	134	20	21	22	22	22	1.9	0.3	0.3	0.3	0.3	0.
	1 0 1 0 0	1150	282	296	302	302	302	16.8	4.1	4.3	4.4	4.4	4.
	1 1 0 0 0	60	79	32	7	3	1	0.9	1.1	0.5	0.1	0.0	0.
	Total	1828	552	525	510	506	505	26.6	8.0	7.7	7.4	7.4	7.
94 FCRS plus two other surveys	**1** 0 0 1 1	12	6	6	6	6	6	0.2	0.1	0.1	0.1	0.1	0.
	1 0 1 0 1	122	42	43	43	43	43	1.8	0.6	0.6	0.6	0.6	0.
	1 0 1 1 0	51	7	8	9	10	10	0.7	0.1	0.1	0.1	0.1	0.
	1 1 0 0 1	16	6	5	1	0	0	0.2	0.1	0.1	0.0	0	
	1 1 0 1 0	4	0	0	0	0	0	0.1	0	0	0	0	
	1 1 1 0 0	20	6	3	0	0	0	0.3	0.1	0.0	0	0	
	Total	225	67	65	59	59	59	3.3	1.0	0.9	0.9	0.9	0.
94 FCRS plus three other surveys	**1** 0 1 1 1	10	3	3	3	3	3	0.1	0.0	0.0	0.0	0.0	0.
	1 1 1 0 1	1	0	0	0	0	0	0.0	0	0	0	0	
	1 1 1 1 0	1	1	1	1	1	1	0.0	0.0	0.0	0.0	0.0	0.
	Total	12	4	4	4	4	4	0.2	0.1	0.1	0.1	0.1	0.
Total		6864	6864	6864	6864	6864	6864	100.0	100.0	100.0	100.0	100.0	100.

† Note: The **Sampling Configuration** indicates the sampling pattern for the 94FCRS, 93FCRS, 94QAS, 94ALS and 94C example, a 10001 indicates sample units selected for only the 94FCRS and 94CSS.

Table 7. For Stratum 90 and above: The Number of 1994 FCRS Samples by State that were in the 199 Stage of the Redrawing Process.

State	Overlap					
	Initial	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
01: Alabama	7	5	1			
04: Arizona	24	22	18	8	5	3
05: Arkansas		4				
06: California	25	21	15	3	1	
08: Colorado	7	7	3	1	1	
09: Connecticut	1					
12: Florida	9	15	7			
13: Georgia	9	11	9	1		
16: Idaho	7	8	4	2	2	1
17: Illinois	22	22	16	3		
18: Indiana	6	6	5			
19: Iowa	13	11	5			
20: Kansas	5	3	2	1		

Table 8. For Stratum 89 and below: The Number of 1994 FCRS Samples by State that were in the 199
Stage of the Redrawing Process.

State	Overlap					
	Initial	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
01: Alabama	3	3	2			
04: Arizona	3	2	2	1	1	1
05: Arkansas	2	1	1	1	1	
06: California	4	3	3	1	1	
08: Colorado	2	2	1			
09: Connecticut	1	1				
12: Florida	6	8	3			
13: Georgia	1	4	2			
16: Idaho	1	1	1			
17: Illinois	6	8				
19: Iowa	7	8	3			
20: Kansas	3	2				
21: Kentucky	1	2	2			
22: Louisiana	2	2				
23: Maine	4	1				
24: Maryland	1					
25: Massachussetts	1	1				
26: Michigan	6	6	4	1		
27: Minnesota	9	6	2			
28: Mississippi	3	2	1			
29: Missouri	1	1				
30: Montana	2	2				
31: Nebraska	2	1				
34: New Jersey	1					
35: New Mexico	1	1				
37: North Carolina	1	1				
38: North Dakota	3	3				
39: Ohio	2	3	3	1		
41: Oregon	2	2	1	1		
42: Pennsylvania	3	1	1			
44: Rhode Island	3	2	2	2	1	1
45: South Carolina	2					
46: South Dakota	3	3	3			
47: Tennessee	1	1				
48: Texas	5	5	2	1		
50: Vermont	1					
53: Washington	1	3	2			
54: West Virginia	1					
55: Wisconsin	1					
Total	102	92	41	9	4	2
Cumulative Percent Reduction		9.8%	59.8%	91.2%	96.1%	98.0%

Table 9. For Stratum 90 and Above, U.S.: The Estimated 1994 Population
Totals for the Initial, First Stage, Second Stage, Third Stage, Fourth Stage, an
FCRS Samples.

Control Variable†	Pop. Total	90% Lower Bound	90% Upper Bound	Estimated Total						Exact CV of Estimated Total	Estim		
				Initial Sample	Stage 1 Sample	Stage 2 Sample	Stage 3 Sample	Stage 4 Sample	Stage 5 Sample		Initial Sample	Stage 1 Sample	Stage Samp
	(#)	(#)	(#)	(#)	(#)	(#)	(#)	(#)	(#)	(%)	(%)	(%)	(%
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Total Land in Farm	206.27	132.60	279.93	197.51	198.46	196.61	203.02	202.30	202.24	21.64	6.39	6.32	6.3

Table 10. For Stratum 89 and Below, U.S.: The Estimated 1994 Population Totals for the Initial, First Stage, Second Stage, Third Stage, Fourth Stage, an FCRS Samples.

Control Variable†	Pop. Total	90% Lower Bound	90% Upper Bound	Estimated Total						Exact CV of Estimated Total	Esti		
				Initial Sample	Stage 1 Sample	Stage 2 Sample	Stage 3 Sample	Stage 4 Sample	Stage 5 Sample		Initial Sample	Stage 1 Sample	Stag Sam
	(#)	(#)	(#)	(#)	(#)	(#)	(#)	(#)	(#)	(%)	(%)	(%)	(%
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14
Total Land in Farm	1288.57	-2847.42	5424.57	604.84	643.63	643.07	642.75	642.79	642.95	194.53	6.59	7.96	7.
Value of Sales	102.97	101.65	104.29	102.64	102.10	102.14	101.99	102.00	102.00	0.77	0.78	0.78	0.
Total Cropland	265.97	260.14	271.80	264.13	263.54	263.11	263.06	263.11	263.11	1.33	1.45	1.50	1.
On-Farm Grain Storage	772.58	-38.39	1583.55	736.54	733.44	735.02	730.05	730.12	730.06	63.62	6.46	7.01	7.
All Cattle & Calves	560.89	539.41	582.36	555.80	555.39	556.54	555.60	555.55	555.75	2.32	2.49	2.50	2.
Total Hogs & Pigs	348.47	312.46	384.49	352.60	349.42	349.37	350.06	350.19	349.88	6.26	5.49	5.64	5.
All Sheep	74.60	49.12	100.08	100.86	102.11	102.14	102.14	102.14	102.14	20.70	18.04	17.88	17.
Farm Workers Hired	86.02	19.45	152.60	81.75	79.73	79.78	79.69	79.68	79.68	46.91	11.52	11.43	11.

† All missing control values were set to zero before any computations were performed. The units of measure and their estimates are:

$1,000,000$ acres for Total Land;
$1,000,000$ acres for Total Cropland;
$100,000$ head for Cattle & Calves, Hogs & Pigs, and Sheep;
$\$1,000,000,000$ for Farm Value Of Sales;
$10,000,000$ bushels for On-Farm Grain Stor
$10,000$ workers Farm Workers Hired.