

Errata: Summary of the 2003 SSBF Reweighting

As a result of the FRB's review of the 2003 SSBF data, fourteen cases that completed the main interview were found ineligible for the survey because they were majority owned by other business firms instead of individuals and fourteen cases were found to be ineligible due to out of scope industry activity. The 28 ineligible cases and their corresponding size class, urban/rural status, and original final weights are shown in Table 1. As a result, these twenty eight cases were dropped from the sample (reducing the sample from 4268 to 4240) and the sample weights recalculated. The reweighting had a minimal effect on response rates, effective sample sizes, design effects, weight variances, or estimated population and subpopulation totals. This document describes the weighting adjustments, provides some comparisons, and updates the weighting information from the 2003 SSBF Methodology report.¹

The reweighting process involved rerunning all weighting programs following the final screener weighting program (wt8ih-wt10ih). It was not necessary to reweight anything prior to the main interview as there was already an established protocol in place to screen out ineligible firms at the main interview. The 28 additional ineligible cases were assigned a main weight (wt8ih) of zero, exactly like the cases that were identified as ineligible during the main interview; eligibility-adjusted and trimmed final weights were then recalculated for the remaining 4240 cases. After calculating the pre-trim weight (w9tih), design effects and pre-trim weights were analyzed in order to determine if there was any need for weight trimming. It was determined that it would be beneficial to keep the original number of cases trimmed in size class 0-19 (0 cases), class 20-49 (4 cases), and class 100-499 (10 cases) but three additional cases were trimmed in size class 50-99 (15 cases). By trimming three additional cases, the design effect for size class 50-99 was brought down to just a little over 2.0. Outlier weights were trimmed to the largest non-outlier weight in the size class (maximum values in Table 5). Last, the summary program was rerun and new response rates, design effects, effective sample sizes, population totals and subtotals were calculated.

¹ The 2003 Survey of Small Business Finances Methodology Report was written by the National Opinion Research Center, Chicago II. It should be noted that the tables in the 2003 SSBF Methodology Report were constructed using the original weights.

There were minor changes in the response propensities, nonresponse cell assignment, and the nonresponse adjustment, although these changes had little impact on the final estimates. The response propensities changed for each case because the total number of pre- and post- reweighted complete and incomplete cases differed. Therefore, when the data was run through the original logistic regression model, each case had a slightly different response propensity score. The changes in propensity score per case caused the average propensity per nonresponse adjustment cell to change as well. Table 2 shows the similarity in the average propensity score within the pre- and post-reweighted nonresponse adjustment cells. After calculating the response propensities for each case, the cases were sorted by ascending response propensity score within size class. Because some of the response propensities changed, the sort order did not remain the same, changing the composition of nonresponse cells slightly. This in turn gave a slightly different response rate within each nonresponse adjustment cell. The change in the response rate within each nonresponse adjustment cell led to slightly different nonresponse adjustments and nonresponse adjusted weights (wt10ih).

Tables 3-9 below compare the weights, response rates, design effect, effective sample sizes and population totals before and after reweighting. We also compare population estimates (see Table 10) before and after reweighting for three key analysis variables: average total employment, average sales, and the distribution of firms by organizational form.

Reweighting had little impact. As shown in Table 3, the main (52%) and the overall (32%) response rates were virtually unchanged. Tables 3-7 provide updated estimates of the design effects and effective sample size by size class and urban/rural status. The design effects for the first two size classes remained similar and there was a slight reduction in the design effect for the two largest size classes. Table 8 shows the total number of firms in the universe prior to reweighting was 6,333,780 compared to post reweighting total of 6,298,088. The loss of 35,692 firms consists of the 15,898 firms represented by the twenty eight affected cases as well as 19,794 additional firms generated by the increased predicted ineligibility rate among the incompletes.² Table 9

² The 15,898 and the 19,794 firms were calculated after the main eligibility adjustment (wt8ih).

shows the similarity between the pre-reweighted and post-reweighted total number of firms by size class and urban rural status.

Table 10 shows the effect the reweighting had on the three key analysis variables: organizational type, employment, and sales. The reweighting had a negligible effect on the total number of employees, the total sales volume, the average number of employees, the average sales volume, the median number of employees, the median sales volume and the percentage of firms by organizational type.

The reweighting proved to have little impact on the final weights and estimates. The estimated number of firms in the target population fell somewhat, as was expected. Although there was little impact on the final estimates, this exercise proved necessary to correct inaccurate information contained in the 2003 SSBF Methodology Report.

1. The identification code, size class, urban/rural status, and final weight of the 28 cases that were dropped from the sample

Obs	SU_ID	Size class	Urban/Rural	Original Final Weight (wt10ih)
1	20126070	(0-19)	Urban	3528.70
2	20183890	(0-19)	Urban	3344.17
3	20030480	(0-19)	Urban	2270.34
4	20311980	(0-19)	Urban	2263.37
5	20075050	(0-19)	Rural	2222.01
6	20043020	(0-19)	Urban	2136.14
7	20320280	(0-19)	Urban	2036.82
8	20004010	(0-19)	Urban	2003.45
9	20374450	(0-19)	Rural	1781.97
10	20017070	(0-19)	Urban	1646.86
11	20350830	(0-19)	Urban	1584.71
12	20202940	(0-19)	Urban	1530.10
13	20231990	(0-19)	Rural	110.70
14	20155900	(0-19)	Urban	84.76
15	20205230	(20-49)	Urban	1745.94
16	20262110	(20-49)	Rural	825.65
17	20092500	(50-99)	Urban	128.64
18	20280940	(50-99)	Urban	78.68
19	20272910	(100-499)	Urban	670.09
20	20007030	(100-499)	Urban	189.98
21	20068710	(100-499)	Urban	145.83
22	20012590	(100-499)	Urban	134.12
23	20194970	(100-499)	Urban	132.05
24	20265140	(100-499)	Urban	103.44
25	20008150	(100-499)	Urban	95.67
26	20085520	(100-499)	Urban	90.23
27	20245710	(100-499)	Urban	67.71
28	20292270	(100-499)	Rural	17.20

2. Differences in Mean propensity scores by nonresponse adjustment cells prior to weight trimming

Nonresponse Adjustment Cell	Size Class	Sample Size	Mean Response Propensity	Mean Response Propensity	Diff in mean Response Propensity
			Before Reweighting	After Reweighting	
1	(0-19)	442	0.711	0.711	0.000
2	(0-19)	441	0.662	0.661	0.001
3	(0-19)	441	0.637	0.636	0.001
4	(0-19)	441	0.616	0.616	0.000
5	(0-19)	441	0.596	0.595	0.001
6	(0-19)	441	0.575	0.574	0.001
7	(0-19)	441	0.554	0.552	0.002
8	(0-19)	441	0.528	0.526	0.002
9	(0-19)	441	0.475	0.472	0.003
10	(0-19)	441	0.371	0.370	0.001
11	(0-19)	441	0.291	0.290	0.001
12	(20-49)	365	0.636	0.635	0.001
13	(20-49)	364	0.53	0.528	0.002
14	(20-49)	364	0.332	0.331	0.001
15	(50-99)	415	0.618	0.616	0.002
16	(50-99)	416	0.487	0.484	0.003
17	(50-99)	416	0.294	0.292	0.002
18	(100-499)	376	0.582	0.577	0.005
19	(100-499)	377	0.418	0.413	0.005
20	(100-499)	377	0.258	0.253	0.005

3. Comparison of Response Rates and Overall Design Effects

Interview Type	Response rate based on Old Weight	Response rate based on new weight	DEFF based on old weight	DEFF based on new weight
Screener	61.92	61.92	NA	NA
Main	52.36	52.41	NA	NA
Overall	32.42	32.45	1.77	1.77

4. DEFF and Effective Sample Sizes by Size Class before reweighting³

Size Class	Count	Max	Mean	St Dev	CV	DEFF	Effective sample size
(0-19)	2842	13263.88	2035.34	1215.97	0.60	1.36	2094.45
(20-49)	569	4576.39	666.70	802.05	1.20	2.45	232.51
(50-99)	444	1456.09	240.01	274.81	1.14	2.31	192.12
(100-499)	413	845.93	153.59	154.42	1.00	2.01	205.40
Total	4268	13263.88	1484.02	1305.41	0.88	1.77	2406.16

5. DEFF and Effective Sample Sizes by Size Class after reweighting⁴

Size Class	Count	Max	Mean	St Dev	CV	DEFF	Effective sample size
(0-19)	2828	13243.15	2033.93	1217.40	0.60	1.36	2082.08
(20-49)	567	4571.32	665.23	803.01	1.21	2.46	230.76
(50-99)	442	1191.84	240.68	246.26	1.02	2.05	215.93
(100-499)	403	841.01	155.28	152.96	0.99	1.97	204.53
Total	4240	13243.15	1485.40	1305.17	0.88	1.77	2392.70

³ This data in this table can be found in Table 6.33 of the 2003 SSBF Methodology Report

⁴ The data in this table updated the data found in Table 6.33 of the 2003 SSBF Methodology Report

6. DEFF and Effective Sample Sizes by Size Class by Urban/Rural before reweighting

Size Class	Urban/ Rural	Count	Max	Mean	St Dev	CV	DEFF	Effective sample size
(0-19)	Urban	2349	13263.88	2030.29	1153.42	0.57	1.32	1775.85
(0-19)	Rural	493	11073.36	2059.37	1479.22	0.72	1.52	325.21
(20-49)	Urban	453	4576.36	674.41	754.80	1.12	2.25	201.10
(20-49)	Rural	116	4576.39	636.58	967.89	1.52	3.31	35.03
(50-99)	Urban	362	1465.09	249.82	275.89	1.10	2.22	163.09
(50-99)	Rural	82	1456.09	196.75	267.39	1.36	2.84	28.80
(100-499)	Urban	346	845.53	161.10	150.44	0.93	1.87	184.82
(100-499)	Rural	67	845.53	114.86	169.48	1.48	3.18	21.09

7. DEFF and Effective Sample Sizes by Size Class by Urban/Rural after reweighting

Size Class	Urban/ Rural	Count	Max	Mean	St Dev	CV	DEFF	Effective sample size
(0-19)	Urban	2338	13243.15	2027.97	1155.53	0.57	1.32	1764.97
(0-19)	Rural	490	10992.89	2062.35	1478.44	0.72	1.51	323.67
(20-49)	Urban	452	4571.32	672.35	755.58	1.12	2.26	199.74
(20-49)	Rural	115	4571.32	637.23	970.53	1.52	3.32	34.64
(50-99)	Urban	360	1191.84	249.94	243.25	0.97	1.95	184.88
(50-99)	Rural	82	1191.84	200.02	256.66	1.28	2.65	30.98
(100-499)	Urban	337	841.01	162.71	148.65	0.91	1.83	183.68
(100-499)	Rural	66	841.01	117.38	169.50	1.44	3.09	21.39

8. Weighted Total of Firms

Size Class	Old Weight		New Weight	
	Count	Weighted Total of firms	Count	Weighted Total of firms
(0-19)	2842	5784429	2828	5751944
(20-49)	569	379350	567	377185
(50-99)	444	106567	442	106380
(100-499)	413	63434	403	62579
Total	4268	6333780	4240	6298088

9. Weighted Total of Firms by Size Class and Urban Rural status

Size Class	Urban /Rural	Old Weight		New Weight	
		Count	Weighted Total of Firms	Count	Weighted Total of Firms
(0-19)	Urban	2349	4769161	2338	4741391
(0-19)	Rural	493	1015268	490	1010553
(20-49)	Urban	453	305507	452	303903
(20-49)	Rural	116	73843	115	73282
(50-99)	Urban	362	90433	360	89978
(50-99)	Rural	82	16134	82	16402
(100-499)	Urban	346	55739	337	54832
(100-499)	Rural	67	7695	66	7747
Total		4268	6333780	4240	6298088

10. Selected Sample Characteristics Before and After Reweighting⁵

		Old Weight (4240)	New Weight (4240)
Total Firms in the Universe		6,302,811	6,298,088
Percentage of firms by Organizational Type (B3)			
	Corporations	15.70	15.69
	Partnerships	8.61	8.65
	S- corporations	31.09	31.07
	Sole Proprietorship	44.59	44.59
Employment (A_TOTEMP)			
	Total	53,124,590.08	53,255,954.75
	Average	8.46	8.48
	Median	3.00	3.00
Sales (P2)			
	Total	6,709,100,000,000	6,724,200,000,000
	Average	1,071,281	1,074,505
	Median	193,000	193,000

⁵ The data was derived from a subset of the Main_final data set as of August 3, 2005 and the variables for sales and total employment were collapsed