On Calibration and Non-response Adjustment for National Compensation Survey

Michael Sverchkov, Alan H. Dorfman, Lawrence R. Ernst, Thomas G. Moerhle, Steven P. Paben, and Chester H. Ponikowski

Contact Author and Presenter: Michael Sverchkov (Sverchkov.Michael@bls.gov), Office of Survey Methods Research, PSB 1950, U.S. Bureau of Labor Statistics, 2 Massachusetts Ave NE, Washington, DC 20212

1. INTRODUCTION AND SUMMARY Most surveys have to deal with non-response. One of the approaches that is used to compensate for this is the correction of the initial weights of respondents through an adjustment factor. Under this approach, the entire sample is divided into cells of similar units, for example, representatives of a particular area, establishment and occupation. For each usable unit (i.e. a unit that contains all necessary information for the future analysis) its frame sample weight is multiplied by the ratio of weighted estimates of the viable units in the cell to weighted estimates of all the usable units in the same cell (non-response adjustment factor). In other words, assuming that the non-respondents in every cell are of the same type as the respondents, the weights of the first are transferred to the last such that the survey estimates of respondent totals coincides with the survey estimates of population totals.

Another simple approach for dealing with non-response is to use post-stratification (benchmarking to known population subtotals). If population totals are known for some subsets of a population (population cells) then instead of correcting the sampling weights for non-response adjustment factors one can correct the sample weight of a usable unit by a ratio of the known total in the cell to its weighted estimate. The advantage of this kind of adjustment is that the weighted benchmark estimator estimates population totals for cells used in the estimation process exactly and is expected to give more accurate results for other kinds of estimates based on weights. Another advantage of this method is that it allows the use of "fresh" information. Very often surveys use the same sampling frame for years even though the frame changes over time, which can lead to biased estimates. Benchmarking makes a correction of the older weights without additional sampling.

Benchmark estimators, and more generally calibration estimators, satisfy in particular the following set of properties: 1) they are weighted estimators with the weights as similar to the original sampling weights as possible, and 2) they estimate perfectly the "known information".

Unfortunately, population totals are usually known only for large cells with units of different types which does not allow us to consider units in a given cell as "similar." Therefore, in practice, calibration estimators are usually applied to non-response adjusted data. As one can see from the description in section 2.1, the non-response adjustment procedure can be complicated and cause large variances of the estimates. Consequently, it would be very useful to know under what conditions the non-response adjustment procedure can be excluded from the estimation processes without producing sizeable biases. The goal of the present research is to find these conditions. Results of the investigation are applied to the National Compensation Survey data.

2. BACKGROUND ISSUES

2.1 Non-Response Adjustment (NRA) for the National Compensation Survey (NCS) The NCS sample of occupational observations (quotes) is selected by a three-step procedure. Simplifying somewhat:

1) The entire country is divided into primary sampling units (PSU) grouped into area strata; one PSU is selected from each stratum with a probability proportional to the employment in the PSU (some PSUs are selected with certainty).
2) Each selected PSU is divided into secondary sampling units (SSU), defined by the establishments in the PSU, also grouped into "industrial" strata; for each "industrial stratum," conditional on the sampled PSU-s, SSU-s are sampled using systematic probability sampling proportionally to the employment in the SSU.
3) Quotes are selected with probabilities proportional to occupation employment, within the selected establishments.

   The NRA procedure involves unit non-response and other weight adjustments performed at both the establishment and the occupational levels.

Establishment Unit Non-Response Adjustment (ENRA) At the establishment level, six different factors are calculated for the data used in the study. These factors are: size-class non-response adjustment factor,

industry non-response adjustment factor, documentation adjustment factor, merge adjustment factor, subsample adjustment factor, and split adjustment factor. These factors can be combined into two groups.

The first group (documentation adjustment factor, merge adjustment factor, subsample adjustment factor, and split adjustment factor) contains documentation-type factors that depend on some external information, for example, discrepancies between the assigned and the actual collected unit, and other similar corrections. We consider these kinds of adjustments as parts of sampling weights.

The second group of factors is the following. The sample is divided into Establishment Non-Response Cells (ENRC). These cells are defined by size of establishments and by the type of industry. In each ENRC the formula for the ENRA Factor is as follows, ENRAF = (A + B) / A, where A= weighted employment of all usable establishments in the ENRC, B= weighted employment of all viable but not usable establishments in the ENRC.

If for an ENRC the ENRAF is not in the range between 1/4 and 4 then the cell is collapsed with the cell assigned by size-class (and on the second stage of adjustment by Industry) collapse pattern (see Requirements for NCS: Unit Non-response), then the above procedure is repeated while the boundary condition 1/4<ENRAF<4 is violated. The finally obtained ENRC is called the Final ENRC (FENRC). Note that this collapsing procedure depends in general on the sample, i.e. Final ENRC can be different for identical initial ENRCs with different samples.

Occupational Unit Non-Response Adjustment (ONRA) At the occupational level there are two adjustment factors: the overall Major Occupational Group (MOG), size-class non-response adjustment factor, and the overall MOG industry non-response adjustment factor. These factors can be described by the following:

The sample is divided into Occupational Non-Response Cells (ONRC). In each cell, the ONRA Factor is defined by ONRAF = (A + B) / A, where A= weighted employment of all usable quotes in the Final ONRC, B= weighted employment of all viable but not usable quotes in the ONRC.

In turn, if an ONRAF is not in the range between 1/4 and 4 then the cell is collapsed with the cell assigned by the collapse pattern that corresponds to one of the two occupational adjustment factors. The finally obtained ONRC is called Final ONRC (FONRC).

The final formula for NRA weights is as follows, NRAW =ENRAF × ONRAF × W, where W is a sample weight.

2.2 Benchmarking in NCS As we mentioned earlier, the benchmark procedure corrects sampling weights of usable units by a Benchmark Factor (BF), i.e., the ratio of the known total in the cell to its weighted estimate. Two potential sets of cells (post-strata) are considered: one with industry strata and one with industry × size-class strata. The totals for these two sets of cells are available from the Longitudinal Data Base (LDB). BF is calculated similarly to NRAF,

BF = (Known total employment in the cell) / (weighted total employment in the cell)

Note that unlike in NRA we cannot use “Occupational Cells” for benchmark adjustment since the LDB does not contain this information.

As in NRA, we follow the restriction: 1/4 < BF < 4. If this condition is violated for some cell then the cell is collapsed with the one assigned by a collapse pattern. We use collapse patterns similar to those used for ENRA.

Typically, the benchmark adjusted weight is defined as BW = BF × NRAW, but we wish to consider also BW = BF × W, where W is a sample weight .

2.3 Does NRA change Benchmark Estimates? If the cells used for NRA and Benchmarking coincide and non-responses arise randomly within the cells used for Benchmarking then one can prove that a benchmark estimate based on the initial frame sample weights does not differ from one based on NRA weights and thus the NRA step does not influence the estimate and can be omitted. It can be shown that similar results hold for some other calibration estimators, see Sverchkov, Dorfman, Ernst, Moerhle, Paben and Ponikowski (2005).

In NCS data, the cells for NRA and benchmarking are different for two reasons. The first reason is that totals for “occupational” cells are not available (at least now), and the second is, even if the initial cells are the same for NRA and Benchmarking, the final cells (result of collapsing) can be different.

The empirical results of section 3 show that for some occupational groups, NRA on the occupational level can appreciably change the final estimate. Indeed if for each industry we have for some size class in an occupational group higher non-response than for the other size classes, then benchmarking on establishment totals can not generally eliminate the difference between non-responses in different size-

classes on an occupational level. This problem might be bypassed if it is possible to get the required totals on the occupational level.

The result of collapsing cells seems to have a small influence on the final estimate.

3. EMPIRICAL INVESTIGATION

3.1 Los Angeles-Riverside-Orange, CA data The practical goal of our research is to test whether the NRA procedure used for the National Compensation Survey (NCS) affects calibration estimates. We started with Los Angeles-Riverside-Orange, CA data for the year 1998 (we consider only the data for which establishments have at least 50 workers). Then we conducted some additional studies of Columbus, OH data of the years 1998 - 2001.

Recall that the Non-Response Adjustment in NCS is defined by type of industry, size of establishment, and occupation non-response cells. For any cell defined by type of industry × size of establishment, the total employment is known and therefore we can use these totals in the benchmark estimation.

At this moment, the total employment for occupational cells is not available for NCS and therefore, only the cells defined by industry × size class cells or larger can be used for benchmarking. On the other hand, NRA is done in industry × occupation × size-class cells. Therefore, we do not know if our main assumption that non-responses arise randomly within the cells used for benchmarking holds or not. Another reason why this assumption can be incorrect is that on both the NRA and the benchmarking stage we follow the requirement that NRAF and BF have to be between the boundaries of ¼ and 4. Since it is possible that for some cells NRAF can satisfy this requirement when BF cannot (or opposite), the final collapsed patterns for NRA and benchmarking can be different.

We carried out an empirical study to assess how much effect these violations of assumption could have in the NCS. We consider the following three classes of cells that seem to be reasonable to use in benchmark estimation:

1) Cells based on 28 industries.
2) Cells defined by 28 industries × 3-class-size indicator where class sizes are defined by the number of employees in the establishment in the range of 50-100, 101-500 and more than 500 respectively (note that these classes do not coincide with size classes used for NRA). In this case, we have 65 final benchmark cells after the collapsing.
3) Cells based on 28 industries × 5-class-size indicator (the same size classes as were used for NRA: 50-99, 100-249, 250-499, 500-999 and 1000 plus), 89 final benchmark cells after the collapsing.

In addition, we consider

4) Cells based on 28 industries × 5-class-size cells but we exclude the sample units and population cell totals for which NRAF or BF is not in the range ¼ - 4 (*Test Set of Cells*).

This last set of cells is not of practical interest since it uses incomplete data. On the other hand, the construction of this set of cells does not involve any cell collapsing. Therefore, if two benchmark estimators based on this set of cells, one based on adjusted and another based on unadjusted weights (that is excluding ENRAF and ONARF), are markedly different, then one can conclude that this difference is caused by Occupational NRA and NRA cannot be omitted from the estimator.

In our empirical studies we estimate Total Number of Employees in the Occupational Group and Hourly Mean Wage in the Occupational Group for 9 Major Occupational Groups (MOGs): Professional, Technical, Executive, Sales, Administration Support, Production and Repair, Machine Operators, Transportation, Handlers, and Services. The estimate for the Hourly Mean Wage in the group is calculated as a ratio of the estimate for Total Annual Wage to the estimate for Total number of hours worked. (See table 1.)

We found that for the *Test Set of Cells* (option 4 above), the difference between benchmark estimators based on sampling weights and NRA weights are small (less than 3% of the estimate value for Totals and less than 1.5% for Mean Wages). We expect that for the Los Angeles-Riverside-Orange, CA 1998 data NRA can be omitted from the estimation process without loss of efficiency. Indeed, for cells based on 28 industries × 5 size classes we obtained results very similar to the "*Test Set of Cells*" results. For the majority of MOGs and cells based on 28 industries and on 28 industries × 3 size classes the differences between adjusted and unadjusted estimators are of the same rate but we obtained some big differences for Transportation. For this group, the difference for Mean Wage is greater than 3% and the same holds for Totals. This results because the ONRAF differ a great deal by size class within the MOG.

3.2 Columbus, OH data We repeated the same studies for Columbus, OH data 1998. We started with a *Test Set of Cells*. The results are far worse than for Los Angeles-Riverside-Orange, CA 1998 data. For example, the difference between the adjusted and unadjusted estimates for Executives (Totals) and Sales

(Mean Wage) are greater than 8% of the estimate. In order to understand whether these differences are caused by a random noise effect, we repeated the study for Columbus, OH data for the years 1999, 2000 (two different samples where collected for that year) and 2001. (See table 2.) The results for "bad" MOGs, Executives (Totals) and Sales (Wages), years 1998, 1999, 2000, Samples 1 and 2 looks similar. This allows us to expect that the response rates in these groups will be different from the average. In other words, responses in a cell depend on the MOG. To check it we investigated a number of NRA cells with high totals for Executives. We found notable differences between the NRAF for Executives and the NRAF of all other MOGs in some of the cells. We concluded that for the Columbus data, NRA can have a high influence on the benchmark estimator in some MOGs.

3.3 Discussion and Recommendations As described in Section 2, NRA involves 3 major steps that can influence the differences between benchmark estimators based on adjusted and unadjusted weights: NRA on the establishment level, NRA on the occupational level, and the collapsing of cells. In the case of the "*Test Set of Cells*", we have no collapsing. Therefore, the differences that we obtained can be based on establishment and/or occupational NRA. In order to decide whether the establishment non-response is important for benchmark estimation, we repeated the previous studies but instead of sample weights, we used weights adjusted for occupational non-response, and we tested the differences between benchmark estimates based on these weights and on fully adjusted weights. . We found that for Columbus, OH 1998 – 2001 (See table 3.) and for Los Angeles-Riverside-Orange, CA 1998 (results not shown), establishment non-response could be omitted from the estimation process when one uses the benchmark estimator without appreciable impact on the estimates. This is not surprising since when we calibrate our weights by establishment totals, we, in fact, make an establishment non-response adjustment to the known totals.

The last studies allow us to expect for NCS that NRA would have very little influence if one could use totals on both the occupation and the establishment level.

One possible source of totals for occupations is the Occupational Employment Survey (OES), which samples 400 thousand establishments per year, gathering wage and employment information on 449 occupations (but not levels within occupation), state agencies doing the actual collection. We have not explored this possibility in depth, regarding that as a new and more complex phase of study, if it is deemed desirable. The following points suggest we would want to do some more preliminary thinking about this before launching any full scale study:

1) Since the categorization of a particular job into an occupation in OES is done, not as in NCS by the field economist, but by the respondent, there is some worry the assignment to MOG would be inaccurate. For example, accounting clerks might in some establishments get confounded with accountants. Accountants are in MOG B, Executive, Administrative, and Managerial Occupations, while accounting clerks are in MOG D, Administrative Support Occupations, including Clerical. It is not clear how much of an impact such misalignment could have. A study, probably involving cognitive experts, would be necessary to settle this question.

2) There could be logistic problems. It is not clear that updated OES data would be available for calibration in a timely fashion. It has been suggested that updating the OES data using NCS data might be in order, yielding a sort of iterative process. Also, the work involved in bringing the two surveys together might itself be quite labor intensive. On the other hand, once the programming technicalities are settled, calibration would likely be routine.

3) The review process, which is where the real labor takes place, will go on whether or not non-response adjustment is carried out, because of the need to check the effects of the documentation factor. Incorporating non-response adjustment into that adds but marginally to the cost in time or effort.

4) There is a question whether, even if the qualms raised in (1) were settled, the OES calibrated estimators, non-response adjusted or not, might be more accurate than estimators calibrated only with respect to establishment totals. The impact of non-ignorable probabilities of response might be magnified under occupational calibration. For example, if a subset of a MOG with lower salaries (relative to the mean MOG salary) is more likely to make it into NCS, but we expand this according to the full number of those in the MOG as estimated from the OES, then the contribution of this MOG to the overall estimator might unduly lower the estimate of wage more than would be the case without calibration.

References

Michael Sverchkov, Alan Dorfman, Lawrence Ernst, Thomas Moerhle, Steven Paben and Chester Ponikowski (2005), On Non-Response Adjustment via Calibration, *2004 JSM Meetings, Proceedings of the Section on Survey Methods Research*

Table 1. Relative differences between benchmark estimators based on unadjusted and adjusted weights. Los Angeles-Riverside-Orange, CA, 1998.

Total Workers by Occupational Group

Occupational Group	***Test Set of Cells***	Cells based on **28 industries x 5-size-classes**	Cells based on **28 industries x 3-size-classes**	Cells based on **28 industry**
All civilian workers	0.00	0.00	0.0	0.0
Executive	-1.67	-1.46	-0.47	-0.76
Sales	-2.59	-0.85	0.90	0.41
Administrative	1.46	2.14	2.18	2.09
Production and Repair	0.53	-1.24	-0.06	1.19
Machine Operators	1.69	1.89	1.79	1.16
Transportation	1.91	2.74	2.40	1.12
Handlers	0.98	1.78	-0.62	-1.09
Services	-0.65	-2.30	-3.13	-2.42
Professional	-0.67	-0.58	0.69	-0.67
Technical	0.78	0.80	1.53	0.64

Hourly Mean Wage by Occupational Group

Occupational Group	***Test Set of Cells***	Cells based on **28 industries x 5-size-classes**	Cells based on **28 industries x 3-size-classes**	Cells based on **28 industries**
All civilian workers	-0.23	-0.35	-0.54	-0.82
Executive	-0.06	-0.40	-0.91	-1.11
Sales	1.25	1.08	1.20	2.93
Administrative	-0.10	-0.06	-0.10	-0.21
Production and Repair	-0.01	-0.24	0.03	-0.20
Machine Operators	0.08	0.06	0.05	0.23
Transportation	-0.26	-1.47	-3.16	-4.64
Handlers	0.10	-0.08	-0.26	-0.83
Services	-0.06	0.03	0.58	0.94
Professional	0.04	0.17	0.17	0.83
Technical	0.37	-0.84	-2.48	-2.85

Table 2. Relative differences between benchmark estimators based on unadjusted and adjusted weights. Columbus, OH, 1998-2001 (*Test Set of Cells*).

Total Workers by Occupational Group

Occupational Group	1998	1999	2000 (1-st sample)	2000 (2-nd sample)	2001
All civilian workers	0.00	0.00	0.00	0.00	0.00
Executive	-8.12	1.38	-2.09	-2.71	7.01
Sales	-0.48	-1.28	-1.22	-2.49	0.21
Administrative	1.14	-0.81	0.04	0.78	0.00
Production and Repair	1.40	-1.25	-4.21	-2.78	1.05
Machine Operators	-0.39	1.77	0.56	0.11	-0.64
Transportation	1.60	2.89	-2.23	-4.52	-0.70
Handlers	-0.10	-1.28	-1.85	-1.05	-6.10
Services	4.08	5.40	4.09	2.29	-6.30
Professional	-1.13	-1.94	2.08	0.87	2.23
Technical	0.64	-0.10	-9.79	0.38	7.68

Hourly Mean Wage by Occupational Group

Occupational Group	1998	1999	2000 (1-st sample)	2000 (2-nd sample)	2001
All civilian workers	-1.72	-0.47	-0.75	-0.85	1.42
Executive	-0.97	1.38	-0.34	-1.15	-0.12
Sales	-8.05	-7.33	-4.90	-6.86	0.11
Administrative	0.11	0.17	-0.27	0.02	0.10
Production and Repair	0.12	0.07	-0.80	-1.55	-0.85
Machine Operators	1.00	1.37	-0.09	-0.35	-0.32
Transportation	-0.69	0.54	0.23	1.50	0.53
Handlers	-0.93	0.00	0.46	-0.13	-0.06
Services	-0.10	0.61	-0.36	-0.14	0.28
Professional	-0.21	-0.18	-0.08	-0.14	0.75
Technical	0.37	0.50	-2.10	0.38	0.47

Table 3. Relative differences between benchmark estimators based on adjusted for occupational non-response and completely adjusted weights. Columbus, OH, 1998-2001 (*Test Set of Cells*).

Total Workers by Occupational Group

Occupational Group	1998	1999	2000 (1-st sample)	2000 (2-nd sample)	2001
All civilian workers	0.00	0.00	0.00	0.00	0.00
Executive	0.27	0.50	0.39	0.40	0.16
Sales	-0.44	0.52	0.68	3.15	-1.41
Administrative	0.24	0.39	0.1	0.10	1.21
Production and Repair	0.32	0.00	0.69	2.39	2.15
Machine Operators	-0.88	-0.74	-0.28	-0.97	-0.76
Transportation	-0.11	0.00	0.02	0.01	1.28
Handlers	-0.13	-0.52	-0.69	-2.38	0.53
Services	-0.11	-0.28	-0.30	-0.12	-0.97
Professional	0.35	-0.25	0.13	-0.53	-0.88
Technical	0.02	0.00	-0.57	0.29	0.46

Hourly Mean Wage by Occupational Group

Occupational Group	1998	1999	2000 (1-st sample)	2000 (2-nd sample)	2001
All civilian workers	0.01	0.05	0.09	0.08	-0.02
Executive	-0.03	-0.23	0.07	-0.11	-0.23
Sales	-2.32	0.30	0.25	-1.47	0.13
Administrative	0.13	0.24	-0.08	0.04	-0.06
Production and Repair	0.08	0.00	0.06	0.43	0.26
Machine Operators	-0.07	0.26	-0.05	-0.22	0.25
Transportation	-0.02	0.00	0.02	0.02	0.20
Handlers	-0.02	0.07	0.19	1.45	0.89
Services	-0.04	0.05	-0.14	0.17	0.41
Professional	0.15	0.05	0.03	-0.20	-0.77
Technical	0.02	0.00	-0.30	0.13	0.05

* Statistics in the tables are calculated as 100x (U-A)/A, where U is the estimate based on the unadjusted weights and A is an estimate based on completely adjusted weights and both benchmark adjustments use the same collapsing pattern

The opinions expressed in this paper are those of the authors and do not necessarily represent the policies of the Bureau of Labor Statistics