Use of Administrative Data To Explore Effect of Establishment Nonresponse Adjustment on the National Compensation Survey Estimates

Chester H. Ponikowski and Erin E. McNulty, Bureau of Labor Statistics
Postal Square Building, Suite 3160, 2 Massachusetts Ave., NE
Washington, DC 20212

Key Words: Unit nonresponse, bias, weighting cells

I. Introduction

Unit non-response is a well known but undesirable problem in sample surveys including the National Compensation Survey (NCS) Program. In NCS, unit non-response occurs because of refusal or inability of a sample establishment to participate in the survey. In addition non-response may occur because of inability of an interviewer to make contact with a sample establishment within a specified survey data collection cycle. Since non-responding sample establishments' data on employee earnings may be systematically different, that is, larger or smaller on average from responding establishments, there may be bias in the survey estimates due to non-response. Non-response also causes an increase in the variance of survey estimates because the effective sample size is reduced. However, bias is usually considered to be a bigger concern because in the presence of a significant bias a calculated confidence interval will be centered on the wrong value and thus will be misleading.

The goal of adjusting for non-response is to reduce bias due to non-response. Over the years, a number of techniques have been presented in statistical literature for adjusting for unit and item non-response (Sverchkov et al, 2005, Kalton and Kasprzyk, 1982; Rubin, 1978; Platek and Gray, 1978). Sverchkov et al propose adjusting for non-response via calibration. Kalton and Kasprzyk describe several methods and their properties. The most common technique for unit non-response is to adjust sampling weights of responding establishments to account for non-responding establishments within a specified set of weighting classes or cells. The weighting cells are defined by available auxiliary variables. The effectiveness of the weighting adjustment in reducing the bias depends on the auxiliary variables' ability to explain the response propensity and the main study estimates, and to identify the most important domains (Sarndol and Lundstrom, 2005).

This paper explores the effect of non-response adjustment on estimates in the National Compensation Survey. We describe the NCS in Section II; present empirical analysis and results in Section III; and state our conclusion and propose issues for further research in Section IV.

II. Description of the National Compensation Survey

The NCS is an establishment survey of wages and salaries and employer-provided benefits conducted by the Bureau of Labor Statistics (BLS). It is the combination of three previously separate surveys: the Employer Cost Index (ECI), the Employee Benefits Survey (EBS), and the locality wage survey. The ECI publishes national indexes that track quarterly and annual changes in wages and benefit costs and also publishes quarterly cost level information on the cost per hour worked of each component of compensation. The EBS annually publishes the incidence and detailed provisions of selected employee benefit plans. The locality wage survey publishes occupational wages for a sample of localities, census divisions, and for the nation as a whole. In addition to the continued publication of these surveys products, new products linking benefit costs and provisions will be published as part of the NCS. All state and local governments and private sector industries, except for farms and private households, are covered in the survey. All employees are covered except the self-employed.

The Longitudinal Database (LDB) serves as the sampling frame for the NCS survey and was used as the administrative data for this study. The LDB is created from State Unemployment Insurance (UI) files of establishments, which are obtained through the cooperation of the individual state agencies.

The integrated NCS sample consists of five rotating replacement sample panels. Each of the five sample panels will be in sample for five years before being replaced by a new panel selected annually from the most current frame. The NCS sample is selected using a three-stage stratified design with probability proportionate to employment sampling at each stage. The first stage of sample selection is a probability sample of areas; the second stage is a probability sample of establishments within sampled areas; and the third stage is a probability sample of occupations within sampled areas and establishments.

Currently the NCS sample consists of 152 areas based on the Office of Management and Budget

(OMB) 1994 area definitions. In 2003 OMB released a new set of area definitions. The new area definitions define a set of Core Based Statistical Areas (CBSA) and designate the remaining geographical areas as outside CBSA counties. The outside CBSA areas for NCS sampling purposes are usually clusters of adjacent counties, not single counties. The NCS has selected a new sample of areas using the 2003 OMB definitions which will replace the current set of primary sampling units (PSUs) over the next few years. A more detailed description of the NCS sample design is available from the BLS website: www.bls.gov/opub/hom/pdf/homch8.pdf.

The NCS locality wage program collects wage data for a sample of occupations within sampled establishments. During the initial interview or update interview some sample establishments refuse to provide or are unable to provide wage data. This results in establishment or unit non-response. Ignoring the establishment non-response could result in substantial bias in estimates and incorrect variance estimates.

In our study, we used the administrative and NCS private industry sample data from the Chicago Consolidated Metropolitan Statistical Area (CMSA) for 2003. The definition of the Chicago CMSA is provided in the 1997 BLS Handbook of Methods. The administrative data provided us with auxiliary variables as well as data on establishment earnings and employment. The administrative data are available approximately nine months after the reference date for the quarterly data collection. The NCS data provided the sample size allocated to private industry in the Chicago area and distribution of NCS non-respondents among industries and establishment size classes. The non-respondents in the NCS are establishments that do not provide any earnings data. The useable establishments are establishments with earnings data for at least one sampled occupation. The in-scope sample size for private industry in the Chicago survey area in 2003 was 651 establishments.

III. Empirical Analysis and Results

To investigate the effect of establishment non-response adjustment on NCS estimates, we calculated and compared response rates for subgroups of the sample, as defined by the auxiliary variables that form the weighting adjustment cells; calculated average earnings for the entire NCS sample, useable establishments, and non-responding establishments; and conducted a simulation study using administrative data.

To find out whether the NCS has a potential problem with bias due to non-response, we compared response rates among subgroups of the sample. If response rates do not vary among subgroups of the sample and missing data are missing at random, then the sample is usually considered not biased as a result of non-response. The NCS defines non-response cells by industry and size class, so we examined response rates in these categories. The response rates were computed by dividing weighted employment of useable establishments by the sum of useable and refusals. The NCS program uses 110 weighting adjustment cells that are defined by industry-size class cross products of the categories listed in Table 1.

The response rates shown in Table 1 indicate that response rates vary widely among industry groups and to a lesser extent among size classes. For example, food stores, finance, insurance, and real estate (FIRE), business services, and education services establishments have response rates much lower than the overall average response rate of 62 percent, while rates in mining, banking, savings and loans, and nursing homes are higher. Size class response rates range from 58 percent for establishments within the 1000-2499 employee group to 67 percent for establishments in the 2500+ employee group.

Table 1. Establishment Response Rates by Industry and Size Class in Chicago

Characteristic	Response Rate
Industry	
Mining	75
Construction	54
Manufacturing Durables	73
Manufacturing Non-Durables	66
Transportation	69
Communications	78
Electric, Gas, Sanitary Services	67
Wholesale	58
Retail	57
General Merchandise Stores	61
Food Stores	44
FIRE	41
Banking, Savings, and Loans	76
Insurance	44
Services	66
Business Services	49
Health	60
Nursing Homes	77
Hospitals	71
Education Services	50
Elementary & Secondary Educ.	73
Higher Education	59
Size Class	

50 – 99	66
100 – 249	60
250 – 999	60
1000 – 2499	58
2500+	67
Total	**62**

The results in Table 1 confirm that industry and size class variables are important auxiliary variables for forming the weighting adjustment cells in NCS. Also the uneven response rates indicate that the NCS may have a potential problem with bias due to non-response.

In our next step of analysis, we assessed how well NCS adjustments for non-response compensate for data lost to establishment non-response. We matched the NCS sample establishments with units on the administrative data file and extracted their earnings and employment information from the file. The earnings data on the administrative file are available at the establishment level only. We calculated average monthly earnings for the respondents, non-respondents, and total sample. The initial sample weights were used in the calculations of estimates. The total sample estimates simulate estimates that might be produced if NCS had no non-response. The estimates based on respondents simulate estimates that might be obtained using initial sample weights that were adjusted for non-response using a single weighting adjustment cell. In addition, we calculated average earnings for respondents using initial sample weights that were adjusted for non-respondents using current weighting adjustment cells and procedures. Collapsing of cells was done using the NCS collapse pattern when adjustment factor was greater than 4.0 within a cell. These estimates simulate published estimates. The results are presented in Table 2, attached at the end of paper.

The average monthly earnings shown in Table 2 indicate that overall average earnings of non-respondents are about 9.6 percent higher than the total sample ($3,934.45 versus $3,588.22). The overall average earnings of respondents are about 8.3 percent lower than the average earnings of the total sample ($3,291.21 versus $3,588.22). When initial sample weights of responding establishments are adjusted for non-responding establishments using current NCS weighting adjustment cells, the overall average earnings of respondents are about 5.6 percent lower ($3,385.88 versus $3,588.22).

The differences in average earnings of respondents, non-respondents, and the total sample are more pronounced for some industry and establishment size class estimates. For example, for services the average earnings for respondents are 12.4 percent lower ($2,282.65 versus $2,606.35) and for non-respondents 21.4 percent higher ($3,163.69 versus $2,606.35) than the average earnings of the total sample. For the establishment size class of 250-999 workers the average earnings for respondents are 14.6 percent lower ($3,025.70 versus $3,543.90) and for non-respondents the average earnings are 18.0 percent higher ($4,183.19 versus $3,543.90) than average earnings of the total sample. However, the differences in average earnings of respondents and non-respondents by industry and employment size class do not always follow this pattern. Non-respondents in establishments with 2,500 or more employees, for example, have lower average earnings than the respondents and the total sample.

In most cases, when initial sample weights of responding establishments are adjusted for non-responding establishments, the differences in average earnings between respondents and the total sample are smaller; non-response procedures tend to bring the respondents' values closer to the actual, full-sample values. Nevertheless, the adjusted estimates continue to lean in the direction of the respondents' data. The results in Table 2 indicate that the NCS locality wage estimates that are generated using weights adjusted for non-response appear to be affected by non-response, that is, the estimates appear to be biased.

The amount of bias in estimated average earnings cannot be determined from a single sample. To measure the amount of bias in the average earnings estimates, we drew a total of 100 samples of the same size and same industry composition as the original sample. The samples were taken from the same frame as the NCS sample in the study; this frame is also the administrative source of the wages and employment figures summarized in Table 2. For each sample a response set was obtained by using the current NCS sample response rates within each non-response adjustment cell. The non-respondents within each non-response adjustment cell were assigned at random.

We generated two sets of estimates. In the first set we used the initial sample establishment weight, and in the second set we used the initial sample establishment weight adjusted for non-response. The sample weight adjustment was done using the current NCS weight adjustment procedures and cells that have five size classes. In addition, we investigated how estimates and variances are impacted when the five employment size classes are collapsed to two employment size classes and then used in sample weight adjustment for non-response. When the adjustment factor exceeded 4.0 within a cell, the collapsing of cells was done using the NCS collapse pattern

The variances for each sample were computed using balanced repeated replication (BRR)

methodology. For a detail description of the BRR methodology see Wolter (1985).

The formulas used to calculate the amount of bias in average earnings and ratios of bias to standard deviation are as follows:

$$B_d = E(\bar{y}_{dr}) - E(\bar{y}_d) = \bar{Y}_{dr} - \bar{Y}_d$$

$$r_d = |B_d| / \sigma_d$$

where,

B_d is the bias in average earnings for domain d

$E(\bar{y}_{dr})$ is the expected value of average earnings of respondents in domain d over the 100 samples

$E(\bar{y}_d)$ is the expected value of average earnings of the total sample in domain d over the 100 samples

$\bar{Y}_{dr}$ is the average earnings of respondents in domain d

$\bar{Y}_d$ is the average earnings in domain d

r_d is the ratio of the bias to standard deviation in domain d

σ_d is the standard deviation for the average earnings in domain d

The results in Table 3 (attached at the end of paper) indicate that the amount of bias in average monthly earnings is reduced when weights are adjusted to account for non-response. In the total monthly earnings estimate of $4,210.03 there is a negative bias of $25.33 when no weight adjustment is done to account for non-response compared to a negative bias of only $3.22 when weight adjustment is done to account for non-response. This is true even when a smaller number of size classes, that is, two size classes, is used in carrying out the adjustment.

The amount of bias in average monthly earnings varies widely among industry groups. Mining, Construction, Manufacturing, Transportation, Utilities, Wholesale and Retail Trade and Education Services show a negative bias ranging from $140.38 to $19.71, and FIRE, Business Services, and Health Services show a positive bias in a range of $10.71 to $220.72. Five out of the eight industry groups have a negative bias.

The number of size classes used in adjustment of weights for non-response seems to have an impact on bias and variance. However, the impact on bias seems to be very minimal. The overall average earnings are underestimated by $3.22 when five sizes classes are used to adjust for non-response and overestimated by $5.44 when two size classes are used in adjustment for non-response. The variance on overall average earnings is $74,474 when two size classes are used compared to $76,900 when five size classes are used in adjustment for non-response. The mean square error on overall average earnings is $74,504 when two size classes are used compared to $76,919 when five size classes are used. The slight reduction in mean square error indicates that the NCS program may benefit from using fewer than five size classes in the adjustment of weights for non-response.

To determine the effect of bias on the accuracy of estimates, we calculated the ratio, r_d, defined above, for each industry group estimate. Cochran (1953) points out that the effect of bias on accuracy of an estimate is negligible if the bias is less than one tenth of the standard deviation of the estimate. A ratio between 0.1 and 0.2 is considered to have a modest impact on accuracy of an estimate. The calculated ratios are presented in Table 4.

The results in Table 4 show that the effect of bias on the accuracy of selected industry estimates is usually negligible. Mining, Construction, Manufacturing, FIRE, Business Services, and Health Services have ratios that are less than 0.10 when the current five size classes are used in the adjustment of weights for non-response. The ratios for Transportation, Utilities, Wholesale and Retail Trade suggest that bias in these estimates has a modest impact on accuracy. The ratio for Education Services is 0.32 which indicates that the bias in this estimate has somewhat noticeable impact on accuracy of this estimate.

Table 4. Ratio of the Bias to the Standard Deviation by Domain and Number of Size Classes Used in Adjustment for Non-response

Domain	r_d *for 5 size classes*	r_d *for 2 size classes*
Mining/Construction	0.02	0.002
Manufacturing	0.07	0.09
Transportation/Utilities	0.14	0.10
Wholesale/ Retail	0.22	0.13
FIRE	0.09	0.12
Business Services	0.08	0.05
Health Services	0.08	0.01
Education Services	0.32	0.23
Total	0.01	0.02

When two size classes are used in weighting adjustments for non-response, the effect of bias on precision of estimates is usually smaller. The ratios are lower for Mining, Construction, Transportation, Utilities, Wholesale and Retail Trade, Business Services, Health Services, and Education Services. The ratios for Manufacturing and FIRE are only slightly higher 0.07 versus 0.09 and 0.09 versus 0.12, respectively. The ratio for Education Services went down from 0.32 to 0.23. The ratios indicate that it might be advantageous to use a different number of size classes for different industry groups in carrying out weighting adjustments for non-response.

IV. Conclusion and Issues for Further Research

In this study, we have explored the effect of establishment non-response adjustment procedures on NCS estimates. Using data from NCS Chicago survey area, we calculated and compared response rates for the auxiliary variables that are used in forming the weighting adjustment cells. We found that response rates vary by industry group and establishment employment size class, the auxiliary variables.

To determine whether non-response might be biasing survey estimates, we used administrative data to calculate average earnings for responding units, non-responding units, and the entire NCS sample in the area. We noted that the NCS weighting adjustment helps reduce the bias due to non-response; the industry and employment size class are powerful auxiliary variables in treating non-response. However, the NCS program could gain from using fewer size classes in forming weighting adjustment cells in some industries.

We selected 100 samples from the original frame and then calculated the ratio of the bias to the standard deviation to assess the effect of bias on the precision of average monthly earnings estimates. We noted that the effect of bias on the precision of estimates is usually negligible. The industry where bias has a modest impact on industry estimates could likely benefit from different number of size class categories.

To extend this study, we would like to examine data from several other survey areas and time periods. We plan to include localities of different size and with different levels of non-response. We plan to compare the direction and magnitude of the bias across time and across areas. If it turns out that there are some consistent trends, then there may be justification for making a non-response bias adjustment. We would like to perform some evaluation of coverage of confidence intervals. We would also like to investigate whether there are any other auxiliary variables that may be useful in reducing bias due to non-response. In particular, we would like to explore whether using average monthly wage as an auxiliary variable would lend strength to re-weighting procedures. We would like to further explore what size class definitions result in the lowest mean square error of NCS estimates. In addition, we would like to investigate the current criteria used for collapsing weighting adjustment cells. As part of this work we would like to determine whether requiring a minimum number of responding establishments within weighting adjustment cells has an impact on bias and variance of estimates. Also, we would like to explore using both the magnitude of the weight adjustment factor and number of responding units in the criteria for collapsing weighting adjustment cells.

References

BLS Handbook of Methods (Bulletin 2490, April 1997), Washington, D.C.: Bureau of Labor Statistics, 57-67.

Cochran, William G. (1953), *Sampling Techniques,* New York: John Wiley & Sons, Inc.

Ernst, L.R., Guciardo, C., Ponikowski, C.H., and Tehonica, J. (2002), "Sample Allocatio and Selection for the National Compensation Survey," *Proceedings of the Section on Survey Research Methods,* Washington, D.C.: American Statistical Association.

Kalton, G., and Kasprzyk, D. (1982), "Imputing for Missing Survey Responses," *Proceedings of the Survey Research Methods Section,* Washington, D.C.: American Statistical Association, 22-31.

Little, R.J.A., (1986), Missing Data Adjustment in Large Surveys, Journal of Business & Economic Statistics, 6, 287-296.

Oh, H. L. and Scheuren, F. J. (1983), Weighting Adjustment for Unit Nonresponse, In W.G. Madow, I. Olkin and D. B. Rubin (editions), *Incomplete Data in Sample Surveys,* Vol. 2, New York: Academic Press.

Platek, R., and Gray, G.B. (1978), "Nonresponse and Imputation," *Survey Methodology,* 4, 144-177.

Rubin, D.B. (1978), "Multiple Imputations in Sample Surveys--A Phenomenological Bayesian Approach to Nonresponse," *Proceedings of the Section on Survey Research Methods,* Washington, D.C.: American Statistical Association, 20-34.

Rizzo, L., Kalton, G. and Brick, J. M. (1996), A Comparison of Some Weighting Adjustment Methods for Panel Nonresponse, *Survey Methodology,* **22**, 43-53.

Sarndal, C.E. and Lundstrom, S. (2005), *Estimation in Surveys with Nonresponse,* London: John Wiley & Sons, Inc.

Sverchkov, M., Dorfman, A. H., Ernst, L.R., Moehrle, T.G., Paben, S. P., and Ponikowski, C.H. (2005), "On Non-response Adjustment via Calibration," *Proceedings of the Section on Survey Research Methods*, Washington, D.C.: American Statistical Association.

Wolter, Kirk M. (1985), *Introduction to Variance Estimation,* New York: Springer-Verlag, Inc.

Any opinions expressed in this paper are those of the authors and do not constitute policy of the Bureau of Labor Statistics.

Table 2. Average Monthly Earnings for NCS Responding, Non-responding, and Total Sample by Selected Industry and Size Class Domains

Domain	*Total Sample*	*Responding Sample Without Weights Adjusted for Non-response*	*Non-Responding Sample*	*Responding Sample with Weights Adjusted for Non-response*
Construction/ Mining	$ 5,009.94	$ 5,441.59	$ 4,440.27	$ 5,197.06
Fire	$ 5,171.23	$ 4,108.36	$ 5,591.18	$ 4,262.07
Manufacturing	$ 4,193.16	$ 4,079.40	$ 4,376.33	$ 4,060.77
Services	$ 2,606.35	$ 2,282.65	$ 3,163.69	$ 2,279.32
TPU	$ 5,322.25	$ 6,227.09	$ 4,255.47	$ 7,214.92
Wholesale/Retail Trade	$ 2,750.03	$ 2,562.55	$ 2,894.30	$ 2,701.75
50-99	$ 3,250.27	$ 3,076.22	$ 3,558.65	$ 3,644.35
100-249	$ 3,267.87	$ 2,692.19	$ 3,727.78	$ 2,614.15
250-999	$ 3,543.90	$ 3,025.70	$ 4,183.19	$ 3,096.14
1000-2499	$ 3,940.40	$ 3,493.42	$ 4,493.30	$ 3,276.69
2500 or more	$ 4,683.79	$ 5,573.08	$ 3,693.30	$ 5,785.78
Total	$ 3,588.22	$ 3,291.21	$ 3,934.45	$ 3,385.88

Table 3. Average Monthly Earnings and Variances for Total Sample and Estimates of Bias Based on 100 Samples by Industry and Number of Size Classes Used in Weights Adjustment for Non-response

Domain	*Total Sample*	*Variance of Total Sample*	*Bias of Responding Sample Without Weights Adjusted for Non-response*	*Variance of Responding Sample*	*Bias of Responding Sample With Weights Adjusted for Non-response*	*Variance of Responding Sample With Weights Adjusted for Non-response*
5 Size Classes						
Construction/Mining	$5,180.73	453,722	$1.08	974,610	-$19.71	908,627
Manufacturing	$5,167.19	105,616	-$50.05	158,501	-$29.56	197,810
Transportation/Utilities	$4,507.31	99,327	$13.22	179,829	-$60.38	191,508
Wholesale/Retail Trade	$2,883.67	53,291	-$5.08	104,979	-$67.89	92,633
FIRE	$9,499.93	2,272,052	-$292.62	3,719,494	$220.72	6,603,822
Business Services	$3,398.98	151,923	$68.29	240,389	$41.93	302,329
Health Services	$3,072.04	14,402	-$18.87	18,339	$10.71	19,178
Education Services	$3,294.37	176,835	-$114.85	190,276	-$140.38	188,517
Total	**$4,210.03**	**31,334**	**-$25.33**	**48,394**	**-$3.22**	**76,909**
2 Size Classes						

Mining/Construction	$5,180.73	453,722	$1.08	974,610	-$1.60	990,795
Manufacturing	$5,167.19	105,616	-$50.05	158,501	-$37.68	183,744
Transportation/Utilities	$4,507.31	99,327	$13.22	179,829	-$41.34	184,630
Wholesale/Retail Trade	$2,883.67	53.291	-$5.08	104,979	-$40.92	96,839
FIRE	$9,499.93	2,272,052	-$292.62	3,719,494	$296.85	6,130,439
Business Services	$3,398.98	151,923	$68.29	240,389	$23.89	279,453
Health Services	$3,072.04	14,402	-$18.87	18,339	-$1.01	20,884
Education Services	$3,294.37	176,835	-$114.85	190,276	-$99.63	190,076
Total	$4,210.03	31,334	-$25.33	48,394	$5.44	74,474