The Weighting Process Used in the Employer Costs for Employee Compensation Series for the National Compensation Survey

Christina L. McCarthy[1], Gwyn R. Ferguson[2], and Chester H. Ponikowski[3]

[1]U.S. Bureau of Labor Statistics, 2 Massachusetts Ave., NE, Room 3160,
Washington, DC 20212

[2]U.S. Bureau of Labor Statistics, 2 Massachusetts Ave., NE, Room 3160,
Washington, DC 20212

[3]U.S. Bureau of Labor Statistics, 2 Massachusetts Ave., NE, Room 3160,
Washington, DC 20212

Abstract

The National Compensation Survey (NCS) is conducted by the Bureau of Labor Statistics (BLS) to compute measures of pay and benefits of America's workers. One output derived from the NCS data is the BLS Employer Costs for Employee Compensation (ECEC) series, which produces quarterly estimates of employer costs for wages, salaries, and benefits. The ECEC weighting process consists of multiple steps including calculation of weights that reflect a 3-stage sampling design, nonresponse adjustment at the establishment level and occupational level, adjustment for sample rotation, and benchmarking. The weighting processes for the survey account for changes in employment totals due to establishments moving out of area, going out of business, or refusing to provide data, employment fluctuations, as well as occupations being abolished or not in-scope for the survey. Data from a recent sample will demonstrate how well the employment totals reflect the in-scope population for the survey throughout each stage of the weighting processes.

Key Words: sampling, sample rotation, nonresponse adjustment, benchmarking

1. Introduction

Survey designs employing multi-stage sampling require a more complex weight derivation process. The weights in multi-stage sample designs reflect each stage of sampling and include several adjustments to the base weights. The base weights are usually defined as the inverse of each unit's probability of selection. In practice, most surveys are voluntary and thus subject to nonresponse. In the presence of nonresponse, survey practitioners are faced with a problem of how to compensate for nonresponse. There are two widely used approaches to compensating for nonresponse: weighting adjustment and imputation. Kalton and Kasprzyk (1986) indicate that weighting adjustments are primarily used to compensate for unit nonresponse while imputation procedures are more likely to be used to compensate for missing items. The NCS utilizes both of these approaches, but this paper focuses on the weighting adjustments.

The NCS is a voluntary establishment based survey with a multi-stage sample design. The NCS sample is subject to sample losses due to establishments being out of business, out-of-scope of the survey, or refusing to participate in the survey at the time of data

collection. Nonresponse in the NCS can occur at the establishment, occupation, or item level. The nonresponse is due to some sample units inability to provide data or due to interviewer inability to contact some sample units or due to respondent refusal. The data from this sample are used to generate several outputs. The weighting process for each output differs slightly. In this paper we describe the weighting process used by the Employer Costs for Employee Compensation (ECEC) series. Section 2 provides an overview of the NCS including survey scope and sample design. Section 3 describes the derivation of initial weights and provides initial weighted sample employment counts. Section 4 describes the weighting adjustment process used to compensate for establishment nonresponse and provides weighted sample employment counts using weights adjusted for nonresponse. Section 5 describes the weighting adjustment process used to compensate for occupational nonresponse and provides weighted employment counts using occupational weights adjusted for nonresponse. Section 6 describes the benchmarking (post-stratification) process used to derive final weights and provides weighted employment counts using final benchmarked weights. Finally, Section 7 provides the concluding remarks.

2. Overview of the NCS

The NCS provides comprehensive measures of employer costs for employee compensation, compensation trends, and the incidence and provisions of employer-provided benefits. A more detailed description of the NCS sample design is given in Chapter 8 of Handbook of Methods and in Ferguson, et al. (2010). The Employment Cost Index (ECI)—a Principal Federal Economic Indicator—is estimated from data collected by the NCS.

The NCS produces several types of data with varying degrees of frequency as summarized below.
- Employment Cost Index (ECI) data are released quarterly
- Employer Costs for Employee Compensation (ECEC) data are released quarterly
- Incidence and Provisions of Employer Provided Benefits data are released annually
- Detailed Provisions for employer provided health insurance, defined benefit retirement plans, and defined contribution retirement plans are released once a year with a focus on one of these benefit areas each year

The NCS covers workers in private industry establishments, and in State and local government, in the 50 States and the District of Columbia. Establishments with one or more workers are included in the survey. Excluded from the survey are workers in the Federal Government and quasi-Federal agencies, military personnel, agricultural industry, workers in private households, the self-employed, volunteers, unpaid workers, individuals receiving long-term disability compensation, individuals working overseas, individuals who set their own pay (for example, proprietors, owners, major stockholders, and partners in unincorporated firms), and those paid token wages.

The BLS Quarterly Census of Employment and Wages (QCEW) serves as the sampling frame for the NCS survey. The QCEW is created from State Unemployment Insurance (UI) files of establishments, which are obtained through the cooperation of the individual state agencies (BLS Handbook of Methods, Chapter 5).

The NCS sample consists of five rotating replacement sample panels for private industry establishments, an additional sample panel for State and local government entities, and an additional panel for private industry firms in the aircraft manufacturing industry. Each of the sample panels is in the sample for at least five years before it is replaced by a new sample panel selected annually from the most current frame. Establishments in each sample panel are initiated over a one-year time period. After initiation, data are updated quarterly for each selected establishment and occupation until the panel in which the establishment was selected is replaced. Estimates for all outputs, except Detailed Provisions, use data from the entire set of 7 independent sample panels.

The NCS sample is selected using a three stage stratified design with probability proportionate to employment size (PPS) 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 selection of occupations (PSO) within sampled areas and establishments. The first stage of the NCS sample consists of 152 areas selected at the national level across geographic areas. These areas or Primary Sampling Units (PSUs) are based on the 2003 Office of Management and Budget (OMB) area definitions.

The second stage of this design occurs at the establishment level within each selected area. Establishments in the sampling frame are stratified by ownership and industry. Industries for the NCS are defined using the North American Industry Classification System (NAICS). Within each of the ownership by industry strata, the NCS employs PPS systematic sampling with frame employment as the measure of size (MOS). To ensure that no unit has a probability of selection greater than one, we identify all units that would be selected with certainty before the sampling process, designate them as part of the sample, and set their sampling weights to one. These certainty units with a weight of one are identified once every five years and are included in each yearly sample until we identify a new set of certainty units. These units are referred to as multi-year certainties. By including them in every annual sample, we ensure that each sample represents the target population while making it easier operationally to process data for the various NCS outputs. After the sample of establishments is selected, it is used for the third stage of the sampling process.

The third stage of this design occurs at the occupational level within each selected establishment. A sample of jobs is drawn from each of these establishments using PPS systematic sampling where the number of employees in the job is the measure of size (see *Appendix B*). To ensure consistency across all establishments, the Standard Occupational Classification (SOC) manual is used to classify the selected jobs into occupations based upon the assigned duties. After this selection and classification we create our smallest aggregate unit known as a quote, which is a distinct combination of time or incentive pay, work level, collective bargaining status, full-time or part-time status, and establishment defined occupation. A more detailed description of the occupational selection is given in Schumann, 2011.

Establishments in each sample are initiated over a one-year time period. During the initiation process, respondents are identified, jobs are selected, and respondents provide the BLS with initial information such as the in-scope employment, type of business and wages for each selected job quote. All establishments are asked to provide the BLS with employer provided wages and salaries for all workers in each selected job quote, the cost of each employer provided benefit, a description of each benefit offered to the employees

in each selected occupation, and benefit access and provisions data such as the number of employees who are offered the benefit, and the number who partake of the benefit.

Respondents are asked to provide quarterly updates for the initiated occupations for the next five years. At the end of the five year update period, the NCS thanks the respondents for supporting our survey and ceases to ask for updated data unless the respondent has been selected in a subsequent sample.

3. Derivation of Initial Weights

We describe the derivation of initial weights for a single private sample which is one of the seven sample panels used in generating quarterly ECEC estimates. This single private sample was selected from the September 2005 QCEW sampling frame. The process of sample selection and weight derivation is the same for each sample panel. The initial weight reflects the sampling of areas and establishments within sampled areas. The initial weight is equal to the inverse of the probability of selecting an area times the inverse of the probability of selecting an establishment within a sampled area. An adjustment factor is then applied to the initial weight to account for a few small sampling strata that had no sample allocated. This adjustment factor reduces the difference between the frame employment and the sample weighted employment at the national level, but it increases the difference between the frame employment and sample weighted employment for some area-industry strata. The **initial weight** formula is as follows:

$$W_{Oi} = \frac{1}{p_a} \times \frac{1}{p_{asi}} \times adj\_fct_{as}$$

Where,

W_{Oi} = is the initial weight of establishment i

p_a = is the inverse of the probability of selecting an area a

p_{asi} = is the inverse of the probability of selecting establishment i in industry s and area a

adj_fct_{as} = adjustment factor in industry s and area a due to allocation

Chart 1 provides for the private industry, the initial weighted employment based on the establishment status from the 1st update after initiation of the sample. For the private industry, the frame employment is approximately 109 million. The initial weighted employment equals the frame employment for the September 2005 sample selection. The data, in the chart below, have been broken out by the update establishment status code. (Please see *Appendix C* for a more detailed explanation of establishment status codes.) The establishments that have no in-scope jobs or dropped employment (employment of establishments with status code of OOB, OOS, or ABO) are not included in nonresponse adjustment and account for approximately 11.2 million in weighted employment. The refusing employment makes up 18.9 million and the usable employment is 78.9 million. Viable establishments are the combination of the usable and the refusal establishments. The viable initial weighted employment is 97.8 million based on the QCEW employment at the time the sample was drawn in September 2005 and is also the NCS in-scope employment using initial weights and frame employment for viable establishments.

Chart 1: Private Industry: Weighted Employment based on Establishment Status and Initial Weights
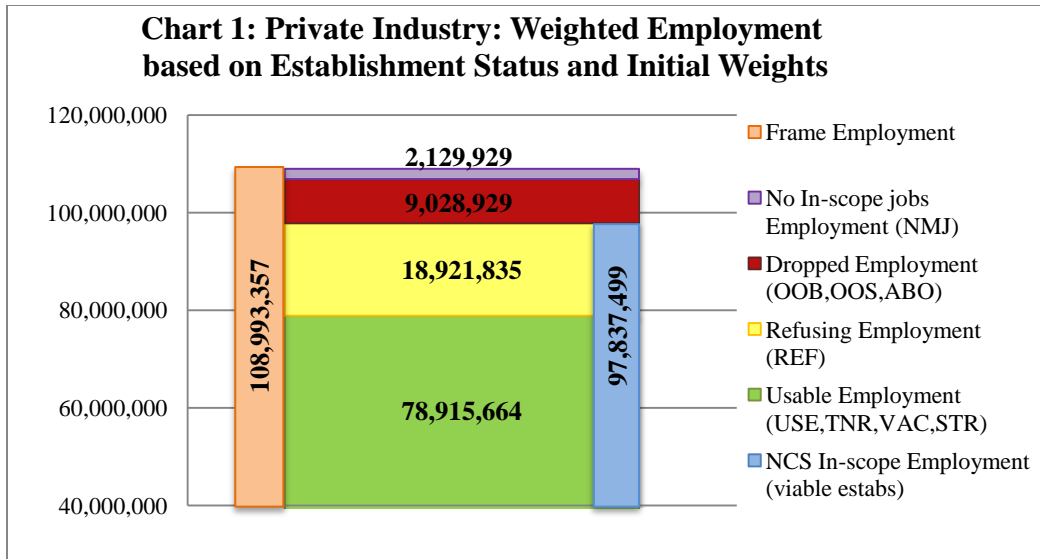
4. Weighting Adjustments for Establishment Nonresponse

Establishment nonresponse adjustment is a process used to redistribute the weights of nonrespondents to similar respondents on the basis of characteristics such as ownership, industry group, and employment size class of the establishment. The process of calculating establishment nonresponse adjustment begins by forming nonresponse cells defined by available auxiliary variables: Ownership (private, State, and local governments), Industry group (defined by NAICS code see **Appendix D**), Size Class (defined by frame employment), and Area. Twenty four industry groups and four size classes are used in forming the cells. Within each formed cell the establishment nonresponse adjustment factor (ENRAF) is calculated using the formula provided below. When a cell ENRAF factor exceeds 4.00 or no usable establishments are in the cell, the system flags the data and then the cells are collapsed using a pre-determined collapse pattern and the factor is re-calculated for the collapsed cell. The collapsing in these situations is done to reduce the impact of nonresponse adjustment on the mean squared error. Establishment nonresponse adjustment is performed after all sample units have been initiated and gone through the first update. Imputation is used to compensate for nonresponse that occurs after initial interview. The establishment nonresponse adjustment is computed at the cell level and applied to the initial weights for all in-scope or cooperating establishments in the cell. The following formula is used to compute the **Establishment NRAF:**

$$ENRAF_{Ci} = \frac{\sum WgtEmp_{CUi} + \sum WgtEmp_{CVi}}{\sum WgtEmp_{CUi}}$$

Where,

$ENRAF_{Ci}$ = is a nonresponse adjustment factor for establishment i in cell C
U_i = in-scope or cooperating establishment i at update for cell C
V_i = refusing to participate establishment i at initiation in cell C
$WgtEmp_{CUi} = FrameEmployment \times W_{Oi}$ = Weighted employment for U_i

$WgtEmp_{CVi} = FrameEmployment \times W_{Oi}$ = Weighted employment for V_i

The nonresponse adjustment process also calculates another factor: the collected other than assigned factor (COTAF). This factor adjusts for any special situations that may

have occurred during data collection and are indicated by a Y or N flag during data collection or validation. There are two possibilities for COTAF, collected more than assigned and collected less than assigned. Collected more than assigned means something more than the original unit was collected. For example, when a sample unit is one of the three establishments owned by a given company and respondent provides data for the three locations instead of the sampled unit, the weight of the sampled unit is adjusted to reflect the employment data actually sampled. Collected less than assigned means that only a piece of the original unit was collected. Computing the **COTAF:**

$$COTAF_i = \frac{Frame\ \ Employment\ \ for\ \ establishment\ \ i}{Collected\ \ Employment\ \ for\ \ establishment\ \ i\ \ at\ \ initiation}$$

(COTAF is set to 1.0 when the flag is set to N)

The nonresponse adjustment process is performed after the update collection for a sample panel. The factors calculated by the system are applied to the initial weight of the appropriate establishment as shown below in computing the **Final Establishment Weight**:

$$FinalEstWgt_i = W_{Oi} \times ENRAF_{Ci} \times COTAF_i$$
(FinalEstWgt$_i$ is set to 0.0 for initiation refusals and all drops)

Where,

$FinalEstWgt_i$ = the Final Establishment Weight

W_{Oi} = Initial Weight of establishment i

$ENRAF_{Ci}$ = Nonresponse Adjustment Factor for establishment i in cell C

$COTAF_i$ = Collected Other Than Assigned Factor of establishment i

The NCS In-scope Employment is 97,837,499, which is the viable weighted employment based on the initial weights and frame employment, from Chart 1 above. The Final Weighted Employment after establishment nonresponse adjustment is 97,838,199. Comparing the weighted employments, we see that the weighted employments are virtually the same confirming that the establishment nonresponse adjustment process reweights the employment appropriately for the refusal units. The weighted employment after ENRA would represent the in-scope employment based on the frame employment from the frame.

## 5. Weighting Adjustments for Occupational Nonresponse

Occupational nonresponse adjustment is a process used to adjust the weights of usable occupations to account for the weights of unusable occupations at initiation. The weight adjustments use a series of factors that, when applied to the final establishment weight of a unit, result in a final occupational weight that is used in the estimation process.

The sample panel occupations are selected and initiated over a one year period and then data updates are conducted each quarter until the sample panel rotates out of the survey (which may be several years). The initiation quarters were collected and nonresponse performed for the occupations at the same time as the establishments. On Chart 2 below, only useable establishments from initiation are available for updating; however, their

status may change at update from the original usable status to that of loss. (Please see *Appendix C* for a more detailed explanation of occupational status codes.)   Like establishments, the longer an occupation remains in the sample, the greater the chance of additional occupational loss due to drops or refusals. Imputation is used to compensate for occupational refusals that occur after initial interview.

Chart 2 compares the private industry in-scope initial weighted employment based on frame employment from Chart 1 (98 million) to the initial weighted employment based on collected employment (95 million). There is a difference of approximately 2.9 million employment (represented in purple). This difference is observed when the frame employment from the September 2005 QCEW is compared to the collected in-scope employment collected at initiation for the viable establishments. The collected in-scope employment differs from the collected and frame employments based on the definitions listed in *Appendix E*. This loss of 2.9 million employment is primarily due to the economic situation at the time of collection.

Chart 2 also contains the private industry final establishment weighted employment by occupational status code from the 1st update after initiation for the sample. The data have been broken out by the update occupational status code. The 1.5 million employment decrease (noted in red) is due to usable establishments having dropped occupations that are not included in nonresponse adjustment. The viable occupations are the combination of usable and viable occupations. The final establishment weighted employment then decreases to 93.4 million which is the in-scope employment for the collected occupations after establishment nonresponse is complete (using occupational status and establishment final weight and collected in-scope employment).

Chart 2. Private Industry: Weighted Employment based on Occupation Status and Final Establishment Weight
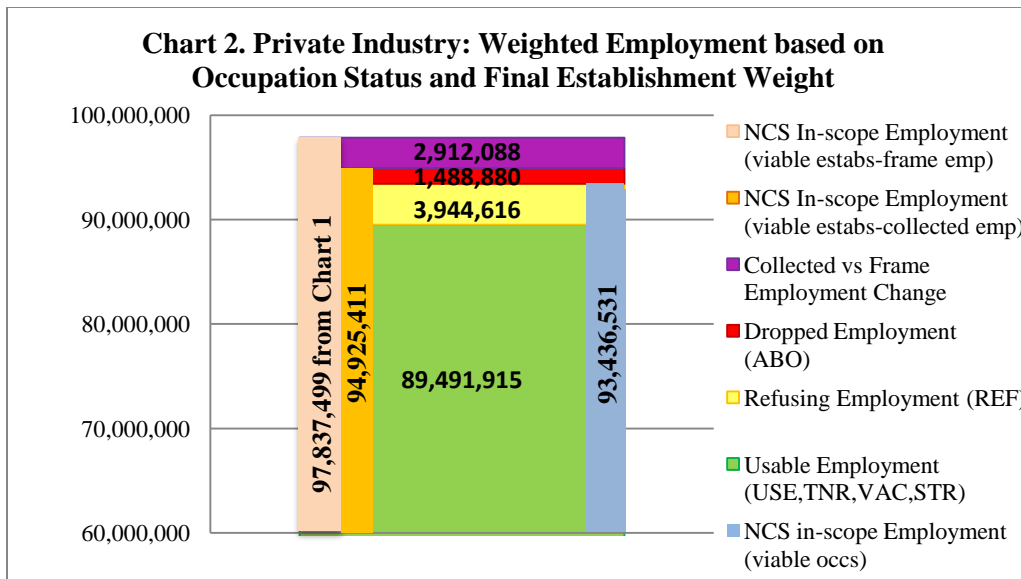
At the occupational level, the nonresponse adjustment is calculated using the occupational nonresponse adjustment factor (ONRAF). The process of calculating ONRAF begins by using nonresponse cells defined at the lowest level, by SOC group *(see Appendix A)*, Industry, Size Class, and Area, and then proceeds to use nonresponse cells formed at the next higher level, if necessary. When a cell ONRAF factor exceeds 4.00 or there are no usable occupations, the system flags the data and then the cells are collapsed using the pre-determined collapse pattern and the factor is re-calculated for the

collapsed cell. The collapsing in these situations is done to reduce the impact of nonresponse adjustment on the mean squared error. Occupational nonresponse adjustment is performed after all sample units have been initiated and gone through the first update. The occupational nonresponse adjustment is computed at the cell level and applied to the final establishment weights for all in-scope or cooperating occupations in the cell:

$$ONRAF_{Cq} = \frac{\sum WgtEmp_{CUq} + \sum WgtEmp_{CVq}}{\sum WgtEmp_{CUq}}$$

Where,

$ONRAF_{Cq}$ = Nonresponse Adjustment factor of quote q in cell C

U_q = in-scope or cooperating quote q at update for cell C

V_q = refusing to participate quote q at update for cell C

$WgtEmp_{C(U,V)q}$ = $PSOInterval *ClpsF_q* FinalEstabWgt_{iq}$ of quote q in cell C

$PSOInterval_{iq}$ = Collected in-scope employment at initiation of establishment i applied
                to quote q / number of assigned quotes *(see Appendix B)*

$ClpsF_q$ = quotes *c*ollapsed during collection (if the same job quote is hit
               more than once during sampling of job quotes within an establishment)

$FinalEstWgt_{iq}$ = the final establishment weight i applied to quote q

After the ONRAF has been calculated, the **final occupational weight** is calculated:

$$FinalOccWgt_q = FinalEstWgt_{iq} \times PSOInterval_{iq} \times ClpsF_q \times ONRAF_{Cq}$$

The Weighted Employment for viable occupations using final establishment weight and the collected in-scope employment at initiation is 93,436,531 million. The weighted employment after occupational nonresponse adjustment using the final occupational weights is 93,436,524 million. These two weighted employments are almost exact showing that the occupational nonresponse process adjusts the weights of the quotes appropriately for the refusing occupations.

6. Benchmarking

Benchmarking or post stratification is the process of adjusting the weight of each occupation in the survey to match the distribution of employment by industry at the reference period. In many large establishment surveys, a considerable amount of time can pass from the time when a sample is selected to when estimates are generated. During this period there is a potential for shifts in employment distribution to occur due to uneven growth or decline in industry employment. Benchmarking helps to correct for these shifts in industry employment and also can offset sampling variability.

In the National Compensation Survey (NCS), the estimation process occurs well after the sample selection process, when economic conditions may have changed. The NCS employs the benchmarking process to create factors that adjust the sampling weight of each sample establishment so that survey estimates are reflective of the economy of the reference period. During the benchmarking process, benchmark factors are calculated for each industry cell. Twenty four private industry cells, ten State government cells, and ten local government cells are used in the benchmarking process (see *Appendix D* for more details).

The first step in the benchmark process is to compute the Sample Panel Factor. This step is performed after the establishment and occupational nonresponse and weight adjustment processes are complete. Since the NCS sample used in estimation consists of multiple sample panels and the weighted employment of each sample panel reflects the entire frame employment, the sample panel reduction factor (SPF) is applied to each sample unit weight so that the sample weighted employment is represented correctly, that is, equals the frame employment. The NCS sample consists of five rotating replacement sample panels for private industry establishments, an additional sample panel for State and local government entities, and an additional panel for private industry firms in the aircraft manufacturing industry. Each of the sample panels is in the sample for at least five years before it is replaced by a new sample panel selected annually from the most current frame. The weights of each sample panel represent the entire frame. Therefore, since benchmarking requires using sample units from more than one sample panel, the sample weights will need to be reduced so that the weighted employment is represented correctly. This reduction of the sample weights is achieved by applying the Sample Panel Factor (SPF) to all weights prior to the estimation process. The calculation is dependent on the specific rotational design of the survey and will change if the design changes. Under the current sample panel design, the sample reduction factor is 0.2 for private industry units, except for aircraft manufacturing and 1.0 for governments units and aircraft manufacturing units. The formulas used in the calculation of the benchmark adjusted occupational weight and benchmark factor are provided below. For this sample panel, the SPF will be 0.2 since it will represent $1/5^{th}$ of the domain. The weighted employment count prior to SPF was 93,436,524 and the weighted employment after SPF was 18,687,305.

Compute the *AdjFinalOccWgt:*

$$AdjFinalOccWgt_q = SPF * FinalOccWgt_q$$

In December 2006, two changes to the ECEC benchmarking process were introduced. The new benchmark process introduced into the ECEC uses data from two BLS programs: the Quarterly Census of Employment and Wages (QCEW) and the Current Employment Statistics (CES) program, as described below. The reason for the change was that the CES is approximately 6 months more current than the QCEW but the industry breaks are not as detailed. Thus, a factor reflecting the change in employment over the six months is applied to the more detailed QCEW employment counts to get the most current employment data available. The use of employment data from the two programs provides the appropriate industry employment coverage and timeliness needed for the ECEC. Secondly, the ECEC benchmark process was refined to account for those establishment/occupations with no in-scope jobs. Typically, such establishments/occupations only include an owner or family members who are out of scope for jobs in our survey. (For more information on ECEC benchmarking, see "Changes in Calculations for the BLS Employer Costs for Employee Compensation Data, March 2007," on the Internet at http://www.bls.gov/ncs/ect/sp/ececcalc.pdf.)

After the sample panel reduction factor has been applied, the **benchmark factor (BMF)** is calculated as follows:

$$BMF_{Cq} = \frac{\sum QCEW\ Employment * (CES_A/CES_B)}{\sum_{USE} AdjFinalOccWgt_q + \sum_{NMJ} (AsgEmp_i * W_{Oi} * SGF_i)} \quad for\ cell\ C$$

Where,

BMF_{Cq} = Benchmark factor for occupation q in cell C

$QCEW\ Employment$ = Most Current Employment available for cell *C*

CES_A = Most Current CES Employment data available for cell *C*

CES_B = CES Employment data from same time as QCEW for cell *C*

NMJ = establishments where all occupations not in scope of survey

When the benchmark factor for an industry cell is larger than 4.00 or when there are fewer than three contributing establishments, the cell is combined or collapsed with one or more other cells to form a new data cell. Cell collapsing is used as a way to minimize variance but with the attempt not to have an adverse effect of increased bias. A maximum factor is capped at 4.00 in order to prevent any one response from being overrepresented in the estimate. Once the BMF is computed, it is applied to all quotes in the cell to calculate the benchmark adjusted occupational weight.

Computing the **Benchmark Adjusted Occupational Weight**:

$BmkAdjOccWgt_q = AdjFinalOccWgt_q * BMF_{Cq}$ for each quote *q* in cell *C*

Table 1 provides the total weighted employment after benchmarking by sample panel. The total benchmarked employment for the private sample panels combined is 106.5 million. The most recent QCEW employment (2nd Quarter of 2008) was 112,866,835 for the benchmarking in February 2009; once the CES factor is applied (4th qtr 2008 CES/2nd qtr 2008 CES by industry) the CES adjusted QCEW Employment, or target population, is 109,665,877.

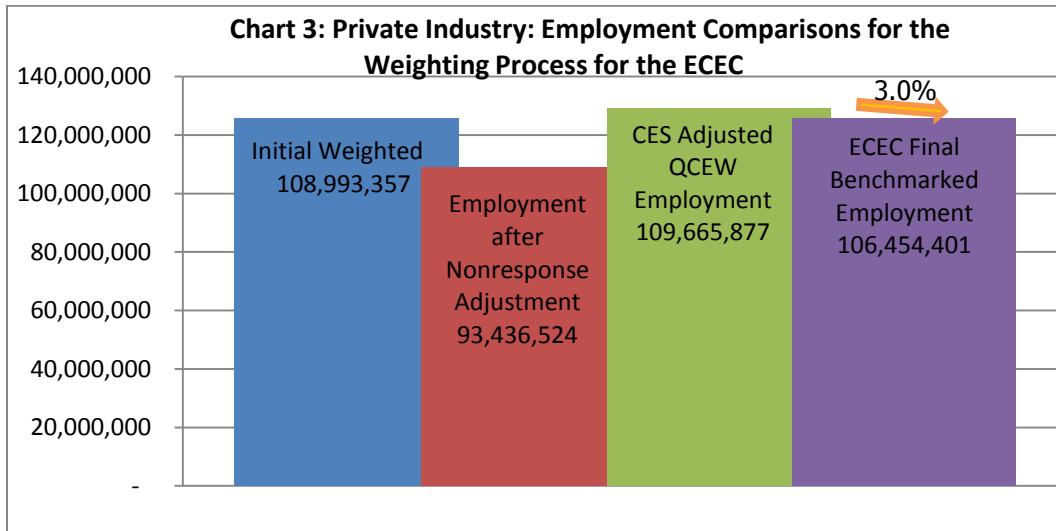**Table 1.** Benchmarked Employment (contains all private sample panels)

Sample Panel (frame)	Weighted Employment after Benchmarking
Private 9/2000	17,649,390
Private 9/2001	19,111,918
Private 9/2002	21,286,338
Private 9/2003	22,709,576
Aircraft manufacturing 9/2004	244,916
Private 9/2005	25,452,263
Total	**106,454,401**

In the benchmark process, there is a current cap of 4.0 on ECEC benchmark factors. However, since the new benchmark procedures have been implemented this process has not been needed since the benchmark factors have not been over 4.0, and typically range from 0.7 to 2.3. The factors are reviewed and validated on a regular basis.

7. Conclusion

The weighting process used in the ECEC series ensures that survey estimates reflect industry-ownership employment counts in proportions consistent with the private industry, State government, and local government sectors at the time of data collection. Chart 3 shows employment counts derived using initial weights, weights adjusted for nonresponse, adjusted QCEW employment, and the benchmarked weights. The private industry frame employment for the ECEC was 108,993,357 from September 2005. After nonresponse adjustment, the ECEC had 93,436,524 weighted employment. The current

employment from the CES adjusted QCEW employment in 2008 was 109,665,877. The final benchmarked employment was 106,454,401. The difference between ECEC total benchmarked employment and the QCEW employment is about 3.0 percent. The difference is due to the discrepancy in worker coverage. Owners and those who set their own pay are examples of establishments/occupations with no in-scope jobs, therefore they are not included in the ECEC, but are included in QCEW. The percent differences between benchmarked estimated employment and CES adjusted QCEW employment are expected to differ somewhat for some subgroups such as industry and/or establishment size classes due to differences in the number of employees in the occupation with no in-scope jobs. The loss of weighted employment due to no in-scope jobs is approximately 3.2 million.

Chart 3: Private Industry: Employment Comparisons for the Weighting Process for the ECEC

Category	Value
Initial Weighted	108,993,357
Employment after Nonresponse Adjustment	93,436,524
CES Adjusted QCEW Employment	109,665,877
ECEC Final Benchmarked Employment	106,454,401

3.0%

References

Kalton, G. and Kasprzyk, D. (1986), "The Treatment of Missing Survey Data", Survey Methodology, Vol. 12, No. 1: 1-16.

Ferguson, G. R., Ponikowski, C. H., and Coleman, J. (2010), "Evaluating Sample Design Issues in the National Compensation Survey", *2010 Proceedings of the Section on Survey Research Methods*, [CD-ROM], Alexandria, VA: American Statistical Association.

Schumann, Richard E., "Occupational Selection and Leveling in the National Compensation Survey", U.S. Bureau of Labor Statistics, Compensation and Working Conditions Online, Originally Posted on August 31, 2011, http://www.bls.gov/opub/cwc/cm20110829ar01p1.htm

Office of Management and Budget (2007). *North American Industry Classification System: United States, 2007.* Washington, D.C.: OMB.

Office of Management and Budget (2007). *Standard Occupational Classification Manual: United States, 2000.* Washington, D.C.: OMB.

U.S. Bureau of Labor Statistics (2008) *BLS Handbook of Methods*, National Compensation Measures, Chapter 8. http://www.bls.gov/opub/hom/homch8_a.htm

U.S. Bureau of Labor Statistics (2008) *BLS Handbook of Methods*, National Compensation Measures, Chapter 5. http://www.bls.gov/opub/hom/homch5_a.htm

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

Appendix A: SOC Aggregate Group (SAG) Cell Definitions

SAG	SOC Code	SOC
1	11, 13	Management, Business, Financial
2	15, 17, 19, 21, 23, 25, 27, 29	Professional and Related
3	31, 33, 35, 37, 39	Service
4	41	Sales and Related
5	43	Office and Administrative Support
6	45, 47	Construction, Extraction, Farming, Fishing
7	49	Installation, Maintenance, and Repair
8	51	Production
9	53	Transportation and Material Moving

Appendix B: Quotes

Private in-scope Emp	Number of Quotes	Government in-scope Emp	Number of Quotes	Aircraft Mfc in-scope Emp	Number of Quotes
1	1	0-49	4	0-49	4
2	2	50-249	6	50-249	6
3	3	250-4,999	8	250-4,999	8
4-49	4	5,000-9,999	10	5,000-9,999	16
50-249	6	10,000-14,999	12	10,000-14,999	32
250+	8	15,000-19,999	14		
		20,000-24,999	16		
		25,000-29,999	18		
		30,000+	20		

Appendix C

Establishment Status Code Definitions
USE-**Usable establishment** – valid status at initiation and update: The schedule contains at least one "USE" occupation. An occupation is classified as "USE" if the following are present: occupational characteristics (FT/PT, U/N, and T/I), work schedule & wage data.
TNR-**Establishment refused temporarily** – usable status at initiation and TNR status at update only: The respondent temporarily would not provide data for any occupation.
VAC-**Vacant occupation** – usable status at initiation and VAC status at update only: All occupations are vacant or other temporary condition (seasonal).
STR-**Establishment on strike** – usable status at initiation and STR status at update: All occupations are on strike.
REF-**Establishment refused to participate** – valid status at initiation or update: The establishment (all occupations) refused to provide any data for any occupation on a permanent basis. No single occupation has earnings, SOC code, worker characteristics, and work schedule.
NMJ-**No in-scope jobs** – valid status at initiation: There are no matching occupations and collected in-scope employment is zero (ex, workers are owners).
OOS-**Establishment out of scope of survey** – valid status at initiation and update: The establishment is in an industry or area outside of the survey's coverage.
OOB-**Establishment out of business** – valid status at initiation and update: The establishment has ceased all productive operations. When an establishment is under new ownership or has a name change it is not out of business.
ABO-**Occupation abolished** – usable status at initiation and ABO status at update: Establishment no longer has any employees in the occupation and does not plan to hire.

Occupational Status Code Definitions
USE-**Usable occupation** – valid status at initiation and update: An occupation is classified as "USE" if occupational characteristics (FT/PT, U/N, and T/I), work schedule and wage data are present for the occupation.
VAC-**Vacant occupation** – usable status at initiation and VAC status at update only: The occupation is temporarily vacant or other temporary condition (seasonal).
STR-**Occupation strike** – usable status at initiation and STR status at update: All occupations are on strike.
REF-**Occupation refused to participate** –valid status at initiation and update: The respondent refused to provide any data for the occupation on a permanent basis. It does not have: earnings, SOC code, worker characteristics, and work schedule.
NMJ-**No in-scope jobs** – valid status initiation: The occupation is selected but is out of scope for PSO. This includes the selection of employees such as owners that should have been excluded from sampling.
ABO-**Occupation abolished** –usable status at initiation and ABO at update: Establishment no longer has any employees' occupation and does not plan to hire.

Status Code Summary:

Status	USE	TNR	VAC	STR	REF	NMJ	OOS	OOB	ABO
Estab.	X	X	X	X	X	X	X	X	X
Occ.	X	X	X	X	X	X			X
Initiation	X				X	X	X	X	
Update	X	X	X	X	X		X	X	X

Appendix D: Industry (NAICS) Cell Definitions

Sampling Cell	NAICS	Industry
PRIVATE		
21A	21	Mining
22A	22	Utilities
23A	23	Construction
31A	31-33(excl 336411)	Manufacturing
31B	336411	Aircraft Manufacturing
42A	42	Wholesale Trade
44A	44-45	Retail Trade
48A	48-49	Transportation & Warehousing
51A	51	Information
52A	52(excl 524)	Finance (excluding Insurance)
52B	524	Insurance Carriers & Related Activities
53A	53	Real Estate and Rental & Leasing
54A	54	Prof, Scientific, & Technical Services
55A	55	Mgmt of Companies & Enterprises
56A	56	Admin. and Support, Waste Mgmt
61A	61(excl 6111-6113)	Education (rest of)
61B	6111	Elementary & Secondary Education
61C	6112,6113	Colleges & Universities
62A	62(excl 622,623)	Health and Social Assistance (rest of)
62B	622	Hospitals
62C	623	Nursing Homes
71A	71	Arts, Entertainment, and Recreation
72A	72	Accommodation and Food Services
81A	81(excl 814)	Other services, except public admin.
LOCAL GOVERNMENT		
10L	21,23,31-33	Goods-Producing
20L	42, 44-45, 48-49, 22	Trade, Transportation, and Utilities
30L	6111	Elementary and Secondary Education
40L	6112,6113	Colleges and Universities
50L	61, excl. 6111-6113	Rest of Education
60L	622	Hospitals
70L	623	Nursing Homes
80L	62, excl. 622-623	Rest of Health and Social Services
90L	92, excl. 928	Public Administration
99L	51-56,71,72,81	Other Service-Producing
STATE GOVERNMENT		
10S	21,23,31-33	Goods-Producing
20S	42, 44-45, 48-49, 22	Trade, Transportation, and Utilities
30S	6111	Elementary and Secondary Education
40S	6112,6113	Colleges and Universities
50S	61, excl. 6111-6113	Rest of Education
60S	622	Hospitals
70S	623	Nursing Homes
80S	62, excl. 622-623	Rest of Health and Social Services
90S	92, excl. 928	Public Administration
99S	51-56,71,72,81	Other Service-Producing

Appendix E: Collected Employment and Collected in-scope Employment
Some workers at an establishment may not be considered in-scope for the purposes of reporting the employment of the establishment. The following chart gives a summary of workers that should be included and excluded from the Reported Employment figure.

Collected Employment

Include	Exclude
All PSO Employees	Outside contractors
Owners	Volunteers
Owner managers	Manufacturer's representative who is self-employed
Proprietors	Employees outside the assigned area
Family members	Individuals on long-term disability (LTD) who are not expected to return
Partners	Non-working individuals with no guarantee to return
Employees who are major stockholder	Employees on strike more than a year
Work/study students	Leased employees
Employees who receive token wages	Individuals on Worker's Compensation who are not expected to return
Temporary Workers	Temporary Help Employees

Some workers at an establishment may not be considered in-scope for the purposes of collected in-scope employment. The following chart gives a summary of workers that should be included and excluded from the collected in-scope employment figure.

Collected in-scope employment

Include	Exclude
Employees on family leave*	Contractors
Employees on executive loan*	Proprietors
Employees on sabbatical*	Owners
Employees on sick leave*	Owner Managers
Employees on strike less than a year*	Major stockholders
Employees on Leave Without Pay (LWOP) during shutdown*	Family members who do not receive market earnings
Seasonal workers	Bona fide partners
Adjunct faculty	Work/study students
Temporary workers/Trainees	Individuals on long-term disability (LTD) not expected to return
Religious workers	Non-working individuals with no guarantee to return
Personalized rate employees	Employees on strike more than a year
Red circle rate employees	Volunteers and unpaid workers
Blue circle rate employees	Employees outside the assigned area
Family members with market earnings	Employees who receive token wages
Substitute teachers	Temporary help employees
Contingent workers	Leased employees
Elected officials/Government funded positions	

*Included if they are taking time off but are expected to return to work in PSO sampling.