Supporting Statement

National Compensation Survey (NCS)

B. Collection of Information Employing Statistical Methods

For detailed technical materials on the sample allocation, selection, and estimation methods as well as other related statistical procedures see the BLS Handbook of Methods, BLS technical reports, and ASA papers listed in the references section. The following is a brief summary of the primary statistical features of the NCS.

Prior planning for the NCS involved the consideration of alternative designs within the overall budgetary constraint. Some of the major elements entering into these considerations were the basic products desired, the availability of data, and the requirements to assure statistically reliable estimates. Other elements considered were the efficiency of alternative collection procedures and the probable degree of cooperation from respondents.

Transition to a redesigned NCS is underway. The final approved 2011 budget called for implementation of an alternative to the LPS component of the NCS, which provided occupational wage data by industry and specific geographic areas. As described in Sections 1 – 3 of this document, the redesigned NCS sample will be selected using a 2-stage stratified design with probability proportional to employment sampling at each stage. The first stage of sample selection is a probability sample of establishments within pre-defined geographic areas of interest, and the second stage of sample selection is a probability sample of jobs within sampled establishments. The NCS will use 24 geographic areas, one for each of the 15 largest metropolitan areas by employment and one for the remainder of each Census Division. Data from all sampled establishments will be used to produce the wage and benefit products.

In FY 2012, NCS will begin selecting samples using the methodology described in this document. Current plans call for collection of the first schedules for the redesigned program in the spring of 2012. To begin transitioning to the redesigned sample BLS has stopped collecting data from the wage-only establishments in the NCS sample. Under the new design, the BLS will revert to a national design in order to preserve the reliability of the ECI and the EBS. With a national design, the BLS will reduce the sample size of the ECI and the EBS by about 25 percent. Section 4.b of this document describes the efforts conducted to develop and test the proposed new sample design. Prior to FY 2012, NCS selected samples using a three-stage sample design that is fully described in Chapter 8 of the BLS Handbook of Methods and in Part B of the 2011 NCS OMB clearance request. As NCS fields the first sample selected using the new design, the Handbook of Methods information will be updated.

1a. Universe

The NCS measures employee compensation in the form of wages and benefits for detailed geographic areas, industries, and occupations as well as national level estimates by industry and occupation. The universe for this survey consists of the Quarterly Contribution Reports (QCR) filed by employers subject to State Unemployment Insurance (UI) laws. The BLS receives these QCR for the Quarterly Census of Employment and Wages (QCEW) Program from the 50 States, the District of Columbia, Puerto Rico, and the U.S. Virgin Islands. The QCEW data, which are compiled for each calendar quarter, provide a comprehensive business name and address file with employment, wage, detailed geography (i.e., county), and industry information at the six-digit North American Industry Classification System (NAICS) level. This information is provided for over nine million business establishments, most of which are in the scope of this survey.

The potential respondent universe used in the selection of the NCS sample of establishments is derived from the QCEW and a supplementary file of railroads for each area in the sample. The QCEW is created from the State Unemployment Insurance (UI) files of establishments, which are obtained through the cooperation of the individual State agencies. UI accounts are assigned to all employers in the United States who are required to pay for unemployment insurance. The NCS universe includes all State and local governments and private sector industries, except for agriculture, forestry, and fishing (NAICS Sector 11) and private households (NAICS Subsector 814). Estimates of the current universe and sample size are about 9,000,000 and 11,400 establishments, respectively. Data on the duties and responsibilities of a sample of jobs will be collected in all sample establishments.

All of NCS's projected 11,400 sample establishments will have quarterly collection for the employment costs and benefits participation and plan provisions.

1b. Sample

Stratification, Sample Allocation, and Sample Selection

Beginning in FY 2012, the NCS sample will be selected using a 2-stage stratified design with probability proportional to employment sampling at each stage. The first stage of sample selection is a probability sample of establishments within pre-defined geographic areas of interest, and the second stage of sample selection is a probability sample of jobs within sampled establishments. For more information on the sample design, see the ASA paper by Ferguson et al titled, "Update on the Evaluation of Sample Design Issues in the National Compensation Survey" (See Attachment A).

Each sample of establishments is drawn by first stratifying the establishment sampling frame by defined geographic area of interest, industry and ownership. The industry strata for private industry are shown below (North American Industry Classification System - NAICS-based). Industry stratification for the State and Local Government Industry is still being studied.

NCS Stratification

Private Industry

Aggregate Industry	Detailed Industry	Included NAICS Codes
Education	Educational Services (Rest of)	61 (excl 6111-6113)
Education	Elementary and Secondary Schools	6111
Education	Junior Colleges, Colleges and Universities	6112, 6113
Finance, Insurance and Real Estate	Finance (Rest of)	52 (excl 524)
Finance, Insurance and Real Estate	Insurance	524
Finance, Insurance and Real Estate	Real Estate, Renting, Leasing	53
Goods Producing	Mining	21
Goods Producing	Construction	23
Goods Producing	Manufacturing (excluding aircraft manufacturing)*	31-33 (excl 336411)
Health Care, including Hospitals and Nursing Care	Healthcare, Social Assistance (Rest of)	62 (excl 622, 623)
Health Care, including Hospitals and Nursing Care	Hospitals	622
Health Care, including Hospitals and Nursing Care	Nursing and Residential Care Facilities	623
Service Providing	Utilities	22
Service Providing	Wholesale Trade	42
Service Providing	Retail Trade	44-45
Service Providing	Transportation and Warehousing	48-49
Service Providing	Information	51
Service Providing	Professional, Scientific, Technical	54
Service Providing	Management of Companies and Enterprises	55
Service Providing	Admin., Support, Waste Management	56
Service Providing	Arts, Entertainment, Recreation	71
Service Providing	Accommodation and Food Services	72
Service Providing	Other Services (excl Public Administration)	81 (excl 814)

*Aircraft Mfg is not included in the overall stratification and allocation of the NCS sample. This sample is handled separately in order to provide estimates to the industry.

After the sample of establishments is drawn, jobs are selected in each sampled establishment. The number of jobs selected in an establishment ranges from 4 to 8 depending on the total number of employees in the establishment, except for government and aircraft manufacturing units and units with less than 4 workers. In governments, the number of jobs selected ranges from 4 to 20. In aircraft manufacturing, the number of jobs selected ranges from 4 for establishments with less than 50 workers to 32 for establishments with 10,000 or more workers. In establishments with less than 4 workers, the number of jobs selected equals the number of workers. The probability of a job being selected is proportionate to its employment within the establishment.

Scope - The NCS sample is selected from the populations as defined above.

Sample Allocation—The total NCS sample will consist of approximately 11,400 establishments. The private portion of this sample, approximately 9,800 establishments, will be allocated based on strata defined by the survey's defined 24 geographic areas and five defined aggregate industries. Details regarding the stratification and allocation of the state and local government sample, consisting of approximately 1,600 establishments, are still being evaluated. Most of the material provided in the remainder of this document involves the survey design as it relates to the private industry. Self-representing, certainty establishments are assigned a sampling weight of 1.00 and non-certainty establishments are assigned a sampling weight equal to the inverse of their selection probability.

For private industry samples, NCS will continue to compute detailed allocations and identify multi-year certainty establishments once every three years under the three-year rotation cycle. If budget or resource levels change significantly, allocation will be recomputed, multi-year certainties will re-selected and a new three-year rotation will begin. This section will first describe the process for allocating sample sizes and identifying the multi-year certainty establishments followed by the proposed methodology for when these processes are not executed for a specified sample.

The sample allocation process starts with a total budgeted sample size. The NCS will use targeted percentages across industries along with the frame data to determine how to distribute the sample units among the sampling cells. Due to the small sample size of the NCS for the private non-aircraft manufacturing allocation, NCS will use a five aggregate-industry stratum allocation with a modified measure of size within each of the 23 detailed industries. This adjustment to the measure of size (MOS) will allow fewer strata but still control the number of units needed in the twenty-three detailed industries for which the NCS wants to publish estimates.

The total three-year NCS private non-aircraft manufacturing sample size is first allocated to the five aggregate-industry strata. The size of each stratum is calculated so that the distribution of the new sample mirrors the desired distribution of the sample in order to maximize the ability to meet publication goals. Next, each of the five aggregate stratum allocations is divided among the 24 geographic areas in proportion to the total adjusted employment of the frame units in the

4

areas, resulting in 120 initial area-industry cell allocations. The MOS of a frame unit is the product of the unit's adjusted employment and an adjustment factor that is used to maintain the current distribution of the sample among the 23 detailed NCS industries.

Multi-year certainty units are identified using the initial cell allocations and the adjusted MOS. Each initial area-industry cell allocation is then reduced by the number of certainty units in the cell to create 120 non-certainty area-industry cell allocations. The MOS adjustment factors are recalculated to exclude the certainty units. Finally, the non-certainty allocations are divided between the three years of the sample design by distributing the non-certainty sample sizes across each of the three years. This distribution is done by dividing each size by three and assigning the integer portion of the result to each of the three years. The remainder is assigned to the appropriate number of years, one establishment at a time, in a manner that allows each annual sample to be the same size and the size for each sampling cell to vary by no more than one from year to year.

Under normal processing, sample allocation and multi-year certainty units will be executed once every three years. During years when prior year allocations and multi-year certainties are being used, the most recently identified set of multi-year certainty establishments will be removed from the frame for operational purposes and added to the final selected sample. This will ensure that each sample group represents the entire frame. The most recent non-certainty allocations and the sample frame without the multi-year certainties will be used to compute the final measure of size adjustment factor and to set the non-certainty sample size for each of the 120 area-industry sampling cells.

Sample Selection - Under this design, NCS will select an independent non-certainty sample of private industry establishments (other than aircraft manufacturing) every year within each of the five aggregate industry and 24 geographic area sampling cells. Within each of the sampling cells, units will be sorted by detailed (23) industry, final adjusted MOS, and establishment identification number. The selection process will follow a systematic Probability Proportionate to Size (PPS) approach where the measure of size includes the adjustment factor as defined in the allocation section. The multi-year certainty units identified in the previous step will be added to each non-certainty sample to form the entire establishment sample each year.

Sample weights will be assigned to each of the selected non-certainty establishments in the sample to represent the non-certainty portion of the frame. Units selected as certainty will be self-representing and will carry a sample weight of one. The sample weight for the non-certainty units will be the inverse of the probability of selection.

2a. Sample Design

Sample Rotation - The NCS plans to move forward with this new national based sample design beginning with the production and collection of the next annual sample. Collection of the first private industry sample under the national based design is scheduled to begin in the spring of 2012. After a 15 month initiation period and one quarter for base period wage and benefit cost collection, this sample will be used in NCS estimates for the first time for December 2013 estimates.

The sample rotation will follow a three year panel design for the private industry sector with a total sample size set for the three year rotation. Large establishments with probabilities of selection greater than their sampling interval based on the three-year sample size will be flagged as multi-year certainty establishments. These units will be selected in the first year and will remain fixed for three years. Independent non-certainty private samples will be selected and initiated every year, with each sample being 1/3 of the total non-certainty sample size.

Although the details have not yet been completed, the plan for the State and Local Government sectors is to select a new single panel sample once every ten years. The first government sector sample will be selected after the initial rotation of three private samples. During the years of sampling and initiating a new State and Local Government sample, the existing private sample will not be replaced. As shown in the chart below, transition to the new design for the private industry will begin in the spring of 2012 with the fielding of the first private industry sample and will continue until late 2015. Private industry estimates computed for December 2013 through September 2015 will be based on data collected under a mix of the old three-stage sample design and the new two-stage sample design.
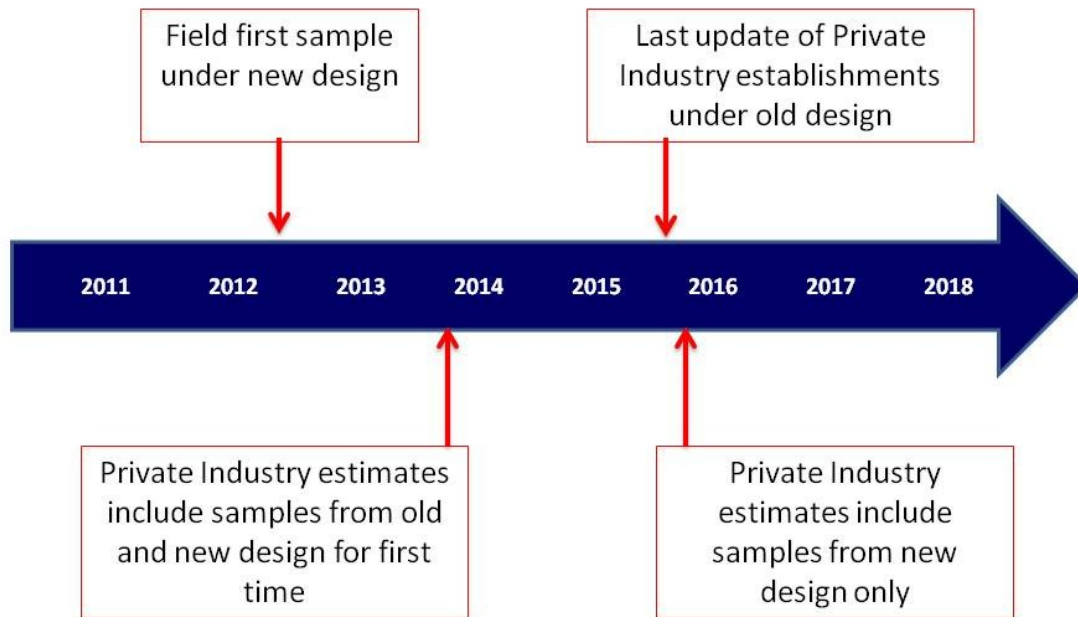
Sample Replacement Scheme

Each year, the NCS selects a new sample of private sector establishments from the most recent available frame data. A new sample of jobs is selected within each private establishment at least once every 3 years (10 years for government) as the establishment is initiated into the survey process. Under this scheme the entire private NCS sample is completely replaced every 3 years.

The primary objectives of the replacement scheme are to reduce the reporting burden of individual establishments by rotating units out of the sample and to insure that the establishment sample is representative of the universe it is designed to cover over time.

A timeline for the NCS sample replacement scheme for the private industry scope is provided below. Details of the transition for the government scope are still being developed.

Proposed Implementation Timeline

Field first sample under new design

Last update of Private Industry establishments under old design

| 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 |

Private Industry estimates include samples from old and new design for first time

Private Industry estimates include samples from new design only

1

2b. Estimation Procedure

The survey produces level estimates, such as average earnings of service workers, and indexes. The estimation procedures for the earnings and index estimates are described below. The Index procedure also includes seasonal adjustment. Note that both of these procedures involve weighting the data from each employee in the sampled job by the final weight.

The final weights include the initial sample weights, adjustments to the initial sample weights, two types of adjustments for non-response, and benchmarking. The initial sample weight for a job in a particular establishment reflects the probability of selecting a particular establishment within one of the pre-defined geographic areas and the probability of selecting a particular job within the selected establishment. Adjustments to the initial weights are done when data are collected for more or less than the sampled establishment. This may be due to establishment mergers, splits, or the inability of respondents to provide the requested data for the sampled establishment. The two types of adjustments for non-response include adjustment for establishment refusal to participate in the survey and adjustment for respondent refusal to provide data for a particular job.

Benchmarking, or post-stratification, is the process of adjusting the weight of each establishment in the survey to match the distribution of employment by industry at the reference period.

Because the sample of establishments used to collect NCS data was chosen over the past several years, establishment weights reflect their employment when selected. For outputs other than the ECI, the benchmark process updates that weight based on current employment. For the ECI, the benchmark process updates that weight based on the employment during the publication base period. For more details about the NCS benchmarking procedures, see the BLS Handbook listed in the references below (Section 6).

The estimation procedure for level estimates uses the benchmark weight, which is the product of the weights, as described in the paragraph above. See Chapter 8 of the BLS Handbook of Methods (available on the BLS Internet at http://www.bls.gov/opub/hom/pdf/homch8.pdf) for an explanation of the estimation procedures for Employer Costs for Employee Compensation estimates and for Benefit Incidence and Provisions estimates.

The index computation involves the standard formula for Laspeyres fixed-employment-weighted index, modified by the special statistical conditions that apply to the NCS. An index for a benefit derived from the NCS data is simply a weighted average of the cumulative average benefit costs changes within each estimation cell, with base-period benefit bills as the fixed weights for each cell. This discussion focuses on the ECI measures of benefit cost changes, but indexes of changes in compensation and wages are computed in essentially the same fashion.

The simplified formula is:

Numerator = $$N = \sum_i W_{0i} M_{ti}$$

Denominator = $$D = \sum_i W_{0i}$$

$$I_t = 100(N/D)$$

where:

i = estimation cell

t = time

I_t = the index at time t

W_{0i} = the estimated base-period benefit bill for the ith estimation cell. The benefit bill is the average benefit cost of workers in the cell times the number of workers represented by the cell.

$$M_{ti} = M_{(t-1)i} R_{ti}$$ = the cumulative average benefit cost change in the ith estimation cell from time 0 (base period) to time t (current quarter).

$M_{(t-1)i}$ = the cumulative average benefit cost change in the ith estimation cell from time 0 (base period) to time t-1 (prior quarter).

R_{ti} is the ratio of the current quarter weighted average benefit cost in the cell to the prior quarter weighted average benefit cost in the cell, both calculated in the current quarter using matched establishment/occupation observations.

The estimation cell is defined on the basis of ownership/industry/major occupation group. For the public sector, separate cells are identified for State and for local governments. Industries as broad as "public administration" and as narrow as "colleges and universities" are treated as separate estimation cell industries. For example, one estimation cell is identified as State government/public administration/clerical workers.

The index computations for the occupation and industry groups follow the same procedures as those for all overall indexes except for the summation. The bills for the occupational groups are summed across industries for each group; the bills for the industry divisions are summed across occupational groups for each industry division.

Computational procedures for the regional, union/non-union, and metropolitan/non-metropolitan measures of change differ from those of the "national" indexes because the current sample is not large enough to hold constant the benefits bills at the level of detail. For these "non-national" series, each quarter the prevailing distribution in the sample between, for example, union and non-union within each industry/occupation cell, is used to apportion the prior quarter benefits bill in that cell between the union and non-union series. The portion of the benefits bill assigned to the union sector is then moved by the percentage change in the union earnings in the cell. The

same is done for the non-union sector. Thus, the relative importance of the union sector in each cell is not held constant over time. Since the relative weights of the region, the union, and the metropolitan area sub-cells are allowed to vary over time, the non-national series are not fixed base period Laspeyres indexes; rather, these are similar to chain linked Laspeyres indexes.

Seasonal Adjustment

Current seasonally adjusted estimates are published in the ECI News Release and historical listing. Each year at the end of the December ECI quarterly production, seasonal adjustment revision is conducted, including revisions to seasonal factors and revisions to historical indexes and 3-month percent changes for the past 5 years. Due to seasonal adjustment revision, the set of published seasonally adjusted series is subject to change each year, as series that are not seasonal are not shown in the seasonally adjusted estimate tables and series that are newly seasonal are added to the tables. Seasonal factors for the coming year are posted on the BLS website at http://www.bls.gov/ncs/ect/ectsfact.htm. Revisions of historical seasonally adjusted data for the most recent five years also appear within the article referenced by the website.

The ECI series are seasonally adjusted using either the direct or indirect seasonal adjustment method. Indexes at comparatively low levels of aggregation, such as the construction wage index, are adjusted by the direct method; that is, dividing the index by its seasonal factor. Seasonal factors are derived using X-12 ARIMA (Auto-Regressive Integrated Moving Average), a seasonal adjustment program developed by the Census Bureau, as an extension of the standard X-11 method. For more information on X-12 ARIMA see the Census website at http://www.census.gov/srd/www/x12a/. Most higher level aggregate indexes are seasonally adjusted by the indirect method, a weighted sum of seasonally adjusted component indexes, where the weights sum to 1.0. For example, the civilian, state and local governments, private industry, goods producing, manufacturing, and service providing series are derived by the indirect seasonal adjustment method.

For more details about the NCS seasonal adjustment procedures, see the BLS Handbook listed in the references below (Section 6).

2c. Reliability

The estimation of sample variances for the NCS survey is accomplished through the method of Balanced Half Samples (BHS). This replication technique uses half samples of the original sample and calculates estimates using those sub samples. The replicate weights in both half-samples are modified using Fay's method of perturbation. The sample variance is calculated by measuring the variability of the estimates made from these sub samples. For a detailed mathematical presentation of this method, see the BLS Handbook of Methods listed in the references.

Before estimates of these characteristics are released to the public, they are first screened to ensure that they do not violate the BLS confidentiality pledge. A promise is made to each private industry respondent and those government sector respondents who request confidentiality, that BLS will not release its reported data to the public in a manner which would allow others to identify the establishment, firm, or enterprise.

Measuring the Quality of the Estimates

The two basic sources of error in the estimates are bias and variance. Bias is the amount by which estimates systematically do not reflect the characteristics of the entire population. Many of the components of bias can be categorized as either response or non-response bias.

Response bias occurs when respondents' answers systematically differ in the same direction from the correct values. For example, this occurs when respondents incorrectly indicate no change in benefits costs when benefits costs actually increased. Another possibility of having response bias is when data are collected for a unit other than the sampled unit. Response bias can be measured by using a re-interview survey. Properly designed and implemented, this can also indicate where improvements are needed and how to make these improvements. The NCS has a Technical Re-interview Program (TRP) that does a records check of a sample of each field economist's schedules of collected data. TRP is a part of the overall review process. TRP verifies directly with respondents a sample of elements originally collected by the field economist. The results are reviewed for adherence to NCS collection procedures. Although not explicitly used to measure bias, this program allows the NCS to identify procedures that are being misunderstood and to make improvements in the NCS Data Collection Manual and training program.

Non-response bias is the amount by which estimates obtained do not properly reflect the characteristics of non-respondents. This bias occurs when non-responding establishments have earnings and benefit levels and movements that are different from those of responding establishments. Non-response bias is being addressed by continuous efforts to reduce the amount of non-response. NCS is analyzing the extent of non-response bias using administrative data from the survey frame. The results from initial analysis are documented in the 2006 ASA Proceedings of Survey Research Methods Section[1]. A follow-up study from 2008 is also listed in the references. Details regarding adjustment for non-response are provided in Section 3 below.

Another source of error in the estimates is sampling variance. Sampling variance is a measure of the fluctuation between estimates from different samples using the same sample design. Sampling variance in the NCS is calculated using a technique called balanced half-sample replication. For national estimates this is done by forming 128 different re-groupings of half of the sample units. For each half-sample, a "replicate" estimate is computed with the same formula as the regular or "full-sample" estimate, except that the final weights are adjusted. If a unit is in the half-sample, its weight is multiplied by *(2-k)*; if not, its weight is multiplied by *k*. For all NCS publications, $k = 0.5$, so the multipliers are 1.5 and 0.5. Sampling variance computed using this approach is the sum of the squared difference between each replicate estimate and the full sample estimate averaged over the number of replicates and adjusted by the factor of $1/(1-k)^2$ to account for the adjustment to the final weights. For more details, see the

---

[1] Ponikowski, Chester H. and McNulty, Erin E., " Use of Administrative Data to Explore Effect of Establishment Nonresponse Adjustment on the National Compensation Survey", 2006 Proceedings of the American Statistical Association, Section on Survey Methods Research [CD-ROM], American Statistical Association, 2006
http://www.bls.gov/ore/abstract/st/st060050.htm

NCS Chapter of the BLS Handbook of Methods. Standard error, which is the square root of variance, for primary aggregate estimates of the index of quarterly change is typically less than 0.5 percent. Relative standard error, which is the square root of variance divided by the estimate, for aggregate estimates of compensation, wage, or benefit levels are typically less than 5 percent. The standard errors or relative standard errors are included within published NCS reports at the following website: http://www.bls.gov/ncs/ect/ectvar.htm.

Variance estimation also serves another purpose. It identifies industries and occupations that contribute substantial portions of the sampling variance. Allocating more sample units to these domains often improves the efficiency of the sample. These variances will be considered in allocation and selection of the future replacement samples.

2d. Data Collection Cycles

NCS data are collected quarterly for all schedules.

3. Non-Response

There are three types of non-response: permanent non-response, temporary non-response, and partial non-response. The non-responses can occur at the establishment level, occupation level, or benefit item level. The assumption for all non-response adjustments is that non-respondents are similar to respondents.

To adjust for permanent establishment or occupation non-response at the initial interview, weights of responding units or occupations that are deemed to be similar are adjusted appropriately. Establishments are considered similar if they are in the same ownership and 2-digit NAICS. If there are no sufficient data at this level, then a broader level of aggregation is considered.

For temporary and partial non-response, a replacement value is imputed based on information provided by establishments with similar characteristics. Imputation is done separately for each benefit both in the initial period and in subsequent update periods. Imputation is also done for each missing wage estimate after the initial period. In the rare event that the BLS cannot determine whether or not a benefit practice exists for a non-respondent, the average cost is imputed based on data from all responding establishments (including those with no plans and plans with zero costs).

There is a continuous effort to maximize response rates. We are developing and providing respondents with new and useful products. Examples include the Program Perspectives Publications (http://www.bls.gov/opub/perspectives/) and industry briefs on selected industries that are provided to field economists to help them identify industry-specific collection challenges. We are continually exploring alternative methods for respondents to report their data. Research is currently underway to provide respondents with additional web-based methods for providing compensation data.

The response rate, based on weighted employment, is expected to be about 74 percent for earnings and benefits initiation schedules. Response rates, based on weighted employment, for

update of earnings only schedules among schedules that responded at initiation is estimated at 93 percent, and 90 percent for earnings and benefits update schedules.

3a. Maximize Response Rates

To maximize the response rate for this survey, field economists initially refine addresses ensuring appropriate contact with the employer. Then, employers are mailed a letter explaining the importance of the survey and the need for voluntary cooperation. The letter also includes the Bureau's pledge of confidentiality. A field economist calls the establishment after the package is sent and attempts to enroll them into the survey. Non-respondents and establishments that are reluctant to participate are re-contacted by a field economist especially trained in refusal aversion and conversion. Additionally, respondents are offered a variety of methods, including telephone, fax, email, and the internet, through which they can provide data.

3b. Non-Response Adjustment

As with other surveys, NCS experiences a certain level of non-response. To adjust for the non-responses, NCS has divided the non-response into two groups, 1) unit non-respondents and 2) item non-respondents. Unit non-respondents are the establishments that do not report any compensation data, whereas item non-respondents are the establishments that report only a portion of the requested compensation data, such as wages for a sub-set of sampled jobs.

The unit non-response is treated using a Non Response Adjustment Factor (NRAF) as explained in the estimation procedure section of this document. Item non-response is adjusted using item imputation. Within each sampling cell, NRAFs are calculated each year based on the ratio of the number of viable establishments to the number of usable respondents in that month. The details regarding the NRAF procedure are given in Chapter 8 of the Bureau of Labor Statistics' Handbook of Methods (http://www.bls.gov/opub/hom/pdf/homch8.pdf).

The method used to adjust for item non-response at the establishment and quote level is a cell-mean-weighted procedure. Details of this procedure are available in BLS Handbook of Methods (http://www.bls.gov/opub/hom/pdf/homch8.pdf). Other techniques are used to impute for item non-response for benefit estimates and are described in the following CWC article: "Recent Modifications of Imputation Methods for National Compensation Survey Benefits Data, may be found at the following link: http://www.bls.gov/opub/cwc/cm20090825ar01p1.htm.

3c. Non-Response Bias Research

Extensive research was done to assess whether the non-respondents to the NCS survey differ systematically in some important respect from the respondents of the survey and would thus bias NCS estimates. Details of this study are described in the two papers by Ponikowski, McNulty, and Crockett referenced in Section 2c (see references below).

4. Testing Procedures

4a. Tests of Collection Procedures

The NCS has developed and is testing a set of new web pages based on its data collection system. Respondent access to these new web pages is occurring through our existing (Internet Data Collection Facility) IDCF system. An initial round of testing was completed with a limited number of respondents during collection of the December 2011 data. In general this test was successful although some minor changes to the web pages were suggested. Based on these results, additional testing will continue with a larger set of respondents for the next few quarters of data collection. Implementation of the new pages will follow several periods of successful testing with respondents.

NCS is also in the initial stages of identifying other potential changes to the data that is collected from respondents, each of which will be tested by the BLS Cognitive Laboratory before full implementation. Some potential changes that could be tested are related to changes in employer provided health care with the implementation of the Affordable Care Act. NCS is currently gathering more information about the potential impact of this legislation on the data collected and provided. Once we better understand the potential impact on our estimates and collected data, NCS will work closely with the Cognitive Laboratory to identify and test the changes to the data collected. Another potential change would be to collect additional details about occupational characteristics that are needed by the Social Security Administration. Any expansion of the data elements desired would be thoroughly tested to ensure that the terms and concepts are understandable and collectible before implementation would begin.

Through environmental scanning, NCS program also identified Payroll Deduction IRA plans as an upcoming compensation trend and started collecting and tracking these plans in 2011. NCS added these plans as a subcategory of the collected Cash or Deferred Arrangements (CODAs) with no employer contributions and is measuring access to this new benefit. This collection contains yes/no/not determinable questions on whether an establishment offers the new benefit. Our research has shown that staff understands the new benefit and its relationship to the existing benefit - Cash or Deferred Arrangements (CODAs) with no employer contributions. This data will be published starting next year.

4b. Tests of Survey Design Procedures

As mentioned previously, the final approved budget for FY 2011 called for an alternative to the LPS component of the NCS, a new approach that uses data from two current BLS programs – the Occupational Employment Statistics (OES) survey and ECI. This may allow for the production of additional locality pay data, while still meeting the requirement to provide data to the President's Pay Agent and to produce the other NCS estimates. With this change the NCS is being redesigned. Transition to a redesigned NCS is underway. These plans call for a transition from an area-based survey design to a non-area-based national design, a reduction in sample size of approximately 25%, and a move from a 5-year rotation cycle to a 3-year rotation cycle. At the same time, NCS has implemented a model-based estimation approach to produce data for the President's Pay Agent. NCS also plans to implement a model-based estimation approach that would allow BLS to continue to produce wage estimates by worker characteristic such as full-time vs. part-time or union vs. non-union. NCS has completed the evaluation and testing for this change in three separate sets of activities.

First, the BLS staff examined potential changes to the NCS sample design that include the following options:

- Moving from an area-based sample design to a national design, thus eliminating the first stage of sampling to select areas
- Implementing a new allocation methodology to correspond with the non-area-based sampling
- Moving from a five-year rotation to a three-year rotation for private industry establishments
- Moving from a design that includes multi-year certainty establishments to a design that controls the number of times each establishment can appear in a 3-year rotation. One option being explored is called dependent sampling.

For each of these options, NCS tested the proposed change using the general scheme described below.

- Obtain a full frame of data,
- Use establishment total wage data from the frame to compute average monthly wages across all establishments,
- Implement the proposed change using the full frame of data,
- Select multiple (100 or more) simulated samples using the proposed methodology,
- Compute estimates of the average monthly wages using the weighted data from each of the simulated samples,
- Compute the mean and standard error of the average monthly wages across all the simulated samples, and
- Compare the estimated average monthly wages across the simulated samples to those from the frame.

In addition to analyzing the potential effect of the redesign on the reliability of the estimates, we also studied the effect of any redesign on response rates and bias as described in the attached paper, "Update on the Evaluation of Sample Design Issues in the National Compensation Survey", by Ferguson et al. Additional work is currently underway to evaluate options for transitioning to a national design for State and local government samples. This research and evaluation is being done using similar methods to those used for the private industry sample design evaluation.

Based on prior experience and a preliminary analysis of the proposed design changes, we believe that the ECI, ECEC, and incidence and key provisions benefits products from the NCS will be of about the same quality as the current estimates. We also believe that we will be able to continue publishing most, if not all, of the current detailed estimates for these product lines. Estimates in the NCS detailed benefits product line are produced from the current initiation sample only. Due to a move to a three-year rotation, each initiation sample will be larger than the current five-year rotation sample even though NCS will implement a sample reduction. The larger sample that will be used to produce the detailed benefits provisions products will hopefully result in some increased accuracy for these estimates, although further evaluation of this is still underway.

Some further work is needed to complete the estimation methodology for the new design and during the transition period. We still need to determine exactly how to create the variance strata used for variance estimation that will be used during the transition. We also need to determine how to create the variance strata to be used once the transition is complete. In all cases, we plan to continue computing variances using the current modified Faye's methodology described in Chapter 8 of the BLS Handbook of Methods. The rest of the estimation processes will use the current methods and formulae. NCS also plans to continue publishing all current outputs for the ECI, ECEC, and NCS Benefits products in the same format and releases as are currently being used.

The BLS staff explored and evaluated different model-based approaches that use data from the OES survey and the ECI portion of the NCS to produce data for the President's Pay Agent. Multiple models were proposed and evaluated using data from recent samples. The proposed models and resulting evaluations are documented in the report, "Using OES Data in Federal Pay Comparability: A Regression-Based Approach" which was shared with the Office of Personnel Management (OPM) during a meeting on April 15, 2009. A copy of this report is attached to this document (see Attachment B).

As described in Attachment B, the BLS introduced a new model-based approach that uses data from two current BLS programs – the OES survey and the ECI, another component of the NCS. In this approach, OES data provides wage data by occupation and by area, while NCS data is used to specify grade level effects. Since the OES sample is much larger than the NCS sample, the BLS expects efficiency gains in the estimates of mean wages by occupation and area. The model-based approach will also be used to extend the estimation of pay gaps to areas not specifically targeted (i.e. the 15 largest MSA's) by the NCS sample. This new approach allowed the BLS to implement an alternative to the LPS component of the NCS, resulting in cost savings.

The model-based approach was presented to the Office of Personnel Management (OPM) in April 2009. During the discussion, OPM indicated that the overall results of this approach appeared to provide high quality data that would meet the requirements of the President's Pay Agent. Moreover, OPM was receptive to the availability of data for additional locality areas. However, OPM did express some concern that the estimates were using data from the OES that are collected without regard to work level, a key component of the pay comparison process since the 1960s and a characteristic specified in FEPCA. The discussion with OPM also highlighted the challenges inherent with presenting this approach to the various stakeholders of Federal locality pay setting.

As documented in Attachment B, BLS has evaluated standard errors associated with the current model used to provide data to the President's Pay Agent and the proposed model using data from NCS and OES. Based on this analysis, for the geographic areas where NCS has data, the proposed regression method appears to be capable of estimating pay gaps with greater precision (lower variance) than does the current approach. However, in studies of small domain estimation over the past thirty years, the predominant practical issues with data quality have involved conditional and unconditional bias, and not variance as such. These bias issues tend to arise from lack of fit in the models employed with these estimators for some domains. Evaluation of these bias issues will require extensive empirical evaluation, and empirical results on bias may vary substantially across time and across the factors used to define the domains of interest. BLS has not completed any bias studies for either the current model or the proposed model and is unable to say whether the new model will change any bias in the modeled estimates.

BLS is also evaluating alternative model-based approaches for using data from the OES survey and the NCS to produce occupational based wage estimates by worker characteristics, such as full-time/part-time status and work level. Although BLS is not mandated to publish specific wage estimates by worker characteristics, appropriate models may be developed that would allow the continued publication of these estimates. The current evaluation of the proposed models includes an analysis of mean-squared error and a comparison of the various models using inclusion probabilities comparing the values produced by various potential estimators to data produced using only NCS data. This work is still in progress.

5. Statistical and Analytical Responsibility

Ms. Gwyn Ferguson, Chief, Statistical Methods Group of the Office of Compensation and Working Conditions, is responsible for the statistical aspects of the NCS program. Ms. Ferguson can be reached on 202-691-6941. As mentioned in the above paragraph, BLS seeks consultation with other outside experts on an as needed basis.

6. References

Bureau of Labor Statistics' Handbook of Methods, Chapter 8, Bureau of Labor Statistics, 2010
http://www.bls.gov/opub/hom/homch8_a.htm#background

Lawrence R. Ernst, Christopher J. Guciardo, Chester H. Ponikowski, and Jason Tehonica, (August 2002), "SAMPLE ALLOCATION AND SELECTION FOR THE NATIONAL COMPENSATION SURVEY," ASA Papers and Proceedings, http://www.bls.gov/osmr/pdf/st020150.pdf

Yoel Izsak, Lawrence R. Ernst, Erin McNulty, Steven P. Paben, Chester H. Ponikowski, Glenn Springer, Jason Tehonica, (August 2005), "UPDATE ON THE REDESIGN OF THE NATIONAL COMPENSATION SURVEY," ASA Papers and Proceedings, http://www.bls.gov/osmr/pdf/st050140.pdf

Chester H. Ponikowski and Erin E. McNulty, (December 2006), "USE OF ADMINISTRATIVE DATA TO EXPLORE EFFECT OF ESTABLISHMENT NONRESPONSE ADJUSTMENT ON THE NATIONAL COMPENSATION SURVEY ESTIMATES," ASA Papers and Proceedings, http://www.bls.gov/osmr/pdf/st060050.pdf

Erin McNulty, Chester H. Ponikowski, Jackson Crockett, (October 2008), "Update on Use of Administrative Data to Explore Effect of Establishment Non-response Adjustment on the National Compensation Survey Estimates," ASA Papers and Proceedings, http://www.bls.gov/osmr/pdf/st080190.pdf

Cochran, William, G., (1977), Sampling Techniques 3rd Ed., New York, Wiley and Sons, 98, 259-261.

Federal Committee on Statistical Methodology, Subcommittee on Disclosure Limitation Methodology, "Statistical Policy Working Paper 22," http://www.fcsm.gov/working-papers/SPWP22_rev.pdf

Matt Dey, Maury Gittleman, Mike Lettau, Steve Miller, (March 2009), "Using OES Data in Federal Pay Comparability: A Regression-Based Approach," (See Attachment B).

Gwyn Ferguson, Chester Ponikowski, Joan Coleman, (August 2010), "Evaluating Sample Design Issues in the National Compensation Survey," ASA Papers and Proceedings, www.**bls.gov/osmr/abstract**/st/st100220.htm

Gwyn Ferguson, Joan L. Coleman, Chester Ponikowski, (August 2011), "Update on the Evaluation of Sample Design Issues in the National Compensation Survey," ASA Papers and Proceedings, (See Attachment A).

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