Supporting Statement

National Compensation Survey (NCS)

**B. Collection of Information Employing Statistical Methods**

**NOTE:**  Proposed Changes to the National Compensation Survey

The President’s proposed budget for FY 2011 calls for an alternative to the Locality Pay Survey (LPS) component of the National Compensation Survey (NCS), a new approach that uses data from two current BLS programs – the Occupational Employment Statistics (OES) survey and Employment Cost Index (ECI) portion of the NCS. 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 continuing to produce the other NCS estimates. At the same time, the NCS will implement a new sample design and implement a sample reduction.

Planning for a redesigned NCS is underway.  These plans, currently being developed, call for a transition from an area-based survey design to a non-area-based national design, as well as other changes that will better reflect the reduced survey scope.  Until the redesign has been fully tested and evaluated, the NCS will continue to select samples using the current design described in Part B.  Section B.4.b outlines the primary technical topics, including important notes on data quality, currently being evaluated for the prospective NCS redesign.  In addition, Ferguson et al. (2010) provide additional technical detail on the prospective redesign research efforts.

Part A of this package reflects the President’s proposed budget in the sample size and respondent burden estimates, outputs, and other aspects of NCS.

Part B of this package reflects the current survey design of the NCS.

**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 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 requirements to assure statistically reliable estimates. Other elements considered were the efficiency of alternative collection procedures and the probable degree of cooperation from respondents.

Current planning for a redesigned NCS is underway. The proposed 2011 budget calls for an alternative to the LPS component of the NCS, which provides occupational wage data by industry and specific geographic areas. As described in Sections 1 – 3 of this document, the NCS sample is selected using a 3-stage stratified design with probability proportional 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 of sample selection is a probability sample of jobs within sampled areas and establishments.  During the selection of establishments, approximately one-half of the establishments, the index portion, are sub-sampled and flagged to support the ECI, ECEC, and NCS Benefits products. The remaining establishments, the wage-only portion, are flagged to support the wage products only. Data from all sampled establishments, both the wage-only portion and the index portion, are used to produce the wage products.

Until the budget change is approved, NCS will continue to select samples using the methodology described in this document. Additionally, NCS will continue to collect data from all establishments selected in the sample to support all the NCS outputs. When the budget is approved to implement the alternative approach, BLS will stop collecting data from the wage-only establishments in the NCS sample. As soon as feasible after the budget implementation, BLS will revert to a national design in order to preserve the reliability of the ECI and EBS. With a national design, the BLS will reduce the sample size of the ECI and EBS by about 25 percent. Section 4.b of this document describes the current efforts to develop and test the proposed new sample design.

**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 eight million business establishments of which about six million 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 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 of all areas in the sample are about 6,000,000 and 31,700 establishments, respectively. Data on the duties and responsibilities of a sample of jobs will be collected in all sample establishments.

Among NCS’s projected 31,700 sample establishments, which include the Locality Pay Survey, approximately 15,400 establishments will have quarterly collection for the employment costs and benefits participation and plan provisions. The approximately 16,300 remaining establishments will have annual collection of earnings data to produce locality and national data.

When NCS implements an alternative to the LPS component of the NCS as described elsewhere in this document, NCS will reduce the size of the sample with quarterly collection by approximately 25% from the current 15,400 establishments to 11,400 establishments. In addition, NCS will stop collecting earnings data from the 16,300 remaining establishments. The burden estimates in Part A reflect these reduced collection efforts.

**1b. Sample**

Stratification and Sample Selection

The NCS sample is selected using a 3-stage stratified design with probability proportional 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 of sample selection is a probability sample of jobs within sampled areas and establishments. For more information on the sample design, see the Bureau of Labor Statistics Handbook of Methods, Chapter 8, at the following website, <http://www.bls.gov/opub/hom/pdf/homch8.pdf>.

The selection of sample areas is done by first dividing the entire area of the United States, consisting of counties (or county equivalents such as parishes in Louisiana) and independent cities, into primary sampling units (PSUs). In most States, a PSU consists of a county or a number of contiguous counties. Metropolitan and micropolitan areas, as defined by OMB, are used as a basis for forming PSUs. Outside of metropolitan and micropolitan areas, a cluster of contiguous counties defines a PSU.

The PSUs with similar average earnings are grouped into strata within each of the 9 Census divisions and three area types (Metropolitan, Micropolitan, outside of metropolitan and micropolitan). One PSU is selected from each stratum with the probability of selection proportional to the employment of the PSU. There are 57 PSUs in strata by themselves that are self-representing, and these include the 27 Combined Statistical Areas (CSAs), the 29 largest Metropolitan Statistical Areas (MSAs), and 1 additional metropolitan area needed to meet the needs of the Pay Agent. The remaining strata are formed by combining PSUs that are MSAs and have similar average annual pay into 60 MSA strata, PSUs that are Micropolitan areas and have similar annual average pay into 22 Micropolitan strata, and PSUs that are outside of metropolitan and micropolitan areas and have similar average annual pay into 13 strata. The PSUs selected with probability proportionate to PSU employment from these strata are non-self-representing because each one chosen represents the entire stratum.

The NCS program started transitioning to the Metropolitan Statistical Areas, Metropolitan Divisions, Micropolitan Statistical Areas, and Combined Statistical Areas in the United States based on the standards published in December 2003, in the Federal Register (65 FR 82228 - 82238) in FY 2007. Current lists of the December 2003 version of the Metropolitan and Micropolitan Statistical Areas and definitions are at this link: <http://www.census.gov/population/www/estimates/metrodef.html>

Each sample of establishments is drawn by first stratifying the sampling frame for each PSU by industry and ownership. The strata for private industry, local government, and State government (North American Industry Classification System -- NAICS based) are as follows:

## NCS Stratification

Private Industry

## NAICS Industry Cell Code Industry 21 21A Mining

22 22A Utilities

23 23A Construction

31-33 (excl 336411) 31A Manufacturing

336411 \* Aircraft Manufacturing

42 42A Wholesale Trade

44-45 44A Retail Trade (rest of)

48-49 48A Transportation and Warehousing

51 51A Information (rest of)

52 (excl 524) 52A Finance (excluding Insurance)

524 52B Insurance Carriers and Related Activities

53 53A Real Estate and Rental and Leasing

54 54A Professional, Scientific, and Tech Services

55 55A Management of Companies and Enterprises

56 56A Admin and Support, Waste Management

61(excl 6111-61130) 61A Education (rest of)

6111 61B Elementary & Secondary Education

6112, 6113 61C Colleges & Universities

62 (excl 622,623) 62A Health and Social Assistance (rest of)

622 62B Hospitals

## 623 62C Nursing Homes

71 71A Arts, Entertainment, and Recreation

72 72A Accommodation and Food Services

81(excl 814) 81A Other services except public administration

\*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.

Local Government

**NAICS Industry Cell Code Industry**

21, 23, 31-33 10L Mining, Constr, Goods-Producing

42, 44-45, 48-49, 22 20L Trade, Transportation, and Utilities

6111 30L Elementary & Secondary Education

6112, 6113 40L Colleges and Universities

61 excl. 6111-6113 50L Rest of Education

622 60L Hospitals

623 70L Nursing Homes

62 excl. 622-623 80L Rest of Health and Social Services

92 excl. 928 90L Public Administration

51, 52-53, 54-56, 71-72, 99L Other Service -producing

81 excl 814

State Government

**NAICS Industry Cell Code Industry**

21, 23, 31-33 10S Mining, Constr, Goods-Producing

42, 44-45, 48-49, 22 20S Trade, Transportation, and Utilities

6111 30S Elementary & Secondary Education

6112, 6113 40S Colleges and Universities

61 excl. 6111-6113 50S Rest of Education

622 60S Hospitals

623 70S Nursing Homes

62 excl. 622-623 80S Rest of Health and Social Services

92 excl. 928 90S Public Administration

51, 52-53, 54-56, 71-72, 99S Other Service –producing

81 excl 814

The number of sample establishments allocated to each stratum is approximately proportional to the stratum employment. Each sampled establishment is selected within a stratum with a probability proportional to its employment. Following the initial allocation and selection of the wage sample, the index, or wage and benefit sample is allocated and selected. The index sample is a subsample of the wage sample. The index sample is roughly half the wage sample. Establishments in the wage sample that are not also included in the index sample are called wage-only units. The details of the allocation process used in the NCS are documented in the 2005 ASA Proceedings of the Survey Research Methods Section[[1]](#footnote-1).

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.

*Stratification*—The NCS sample, including the LPS, has 31,700 establishments allocatedbased on the stratification of ownership (private, state government, local government), nine Census divisions, 24 industries for private, and 10 industries for state and local government. 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. Establishment counts for the sample, by area and ownership, are shown in Appendix A. In-scope private employment and establishment counts for the NCS Survey Areas are summarized by geographic area in Appendix A.

**2a. Sample Design**

*Allocation method*—The current NCS sample, which includes the LPS is a stratified probability-proportional-to-size systematic random sample. The basic sampling unit is an establishment or worksite. Sampled state and local government units, as well as aerospace units, generally remain in the survey for 11 years. Private units remain in the survey for 6 years. The characteristics used to stratify the establishment sample are geographic area by nine Census divisions, and industry divisions defined primarily by 2-digit NAICS (see Section 1b. for strata definitions and Appendix A for establishment allocation).

NCS characteristics are highly correlated with an establishment’s geographic area, industry and employment level. Thus for a fixed sample size, stratified sampling results in a greater precision than simple random sampling. Some establishments are included in the sample with certainty.

## *Sample Rotation*—See the section on Panel Structure in the BLS Handbook of Methods listed in the references (Section 6).

Sample Replacement Scheme

NCS, which currently includes the LPS, selects a new sample of areas approximately once every ten years. The current sample of areas was selected from areas that were defined by OMB after the 2000 Census. In 2004, NCS selected a new sample of areas from areas that were based on the 2000 Census. Each year, NCS selects a new sample of establishments from the most recent available frame data. Private industry establishments from the old area sample are rotated out of the NCS survey outputs over a period of approximately 6 years as the new area private industry establishment samples are rotated into the survey outputs. Beginning in 2006, a new sample of State and local government establishments was selected and was introduced into NCS at the end of 2007. The State and local government establishments will remain in the sample at least 11 years. A new sample of jobs is selected within each private establishment at least once every 6 years (11 years for government) as the establishment is initiated into the survey process. Under this scheme the entire private NCS sample is completely replaced every 6 years.

The primary objectives of the replacement scheme are to reduce 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.

Appendix B provides an overview of the NCS sample replacement scheme.

*Research on sample issues***—**As mentioned in the introduction, proposed NCS 2011 budget levels may require the NCS to move from an area-based design to a non-area-based national design. Research is underway to explore alternative sample designs. See Section 4.b for more details.

**2b. Estimation Procedure**

The survey produces level estimates, such as average earnings of professional workers at the entry level, along with quartiles, first and last deciles, 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 and PSU reflects the probability of selecting a particular PSU, the probability of selecting a particular establishment within the PSU, and the probability of selecting a particular job within the selected establishment and PSU. 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 or 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, such as mean weekly earnings, mean annual wages, and mean hourly earnings, use the individual weight, which is the product of the weights, as described in the paragraph above, of the sampled job, the individual rates in the sampled job, and the number of weeks worked per year. The calculation of the mean hourly earnings includes, in addition to the individual weight, individual wage rate, and number of weeks worked per year, the number of hours paid per week according to the employee’s standard work schedule. For mean weekly earnings this involves multiplying the weekly wage rate for each employee in the sample job by the final weight and the number of annual weeks worked, summing, and dividing by the sum of the final weights times the number of workers for which the NCS collected data times the number of weeks worked. 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 = 

Denominator = 



where:

i = estimation cell

t = time

 is the index at time t

 is 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.

 is the cumulative average benefit cost

change in the ith estimation cell from time 0 (base period) to time t (current quarter).

 is the cumulative average benefit cost change in the ith estimation cell from time 0 (base period) to time t-1 (prior quarter).

 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/nonunion, 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 nonunion within each industry/occupation cell, is used to apportion the prior quarter benefits bill in that cell between the union and nonunion 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, and similarly for the nonunion 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 replicates 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 Bureau of Labor Statistics’ (BLS) confidentiality pledge. A promise is made by the Bureau to each respondent 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. Wage estimates which fail confidentiality screening based on p-percent rule for disclosure (see Federal Committee on Statistical Methodology Working paper 22) are not published. Other statistical procedures are used to determine the disclosure limitations of other estimates.

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[[2]](#footnote-2). A follow-up study from 2008 is also listed in the references. Details, regarding adjustment for nonresponse, 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 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 are 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 index schedules and annually for wage 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 plans to provide industry briefs 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 web-based methods for providing compensation data.

The response rate, based on weighted employment, is expected to be about 78 percent for earnings only initiation schedules and 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, interviewers 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, and pledging confidentiality. An interviewer 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 an interviewer especially trained in refusal aversion and conversion. Additionally, respondents are offered a variety of methods, including telephone, fax, email, and internet, by 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-response. Unit non-respondents are the establishments who do not report any compensation data and item non-respondents are the establishments who report only a portion of the requested compensation data, for example, wages for a subset 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 and 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 the Bureau of Labor Statistics’ Handbook of Methods, Chapter 8 (see <http://www.bls.gov/opub/hom/pdf/homch8.pdf>).

The method used for item imputation for wage estimates 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 imputation techniques are used for benefit estimates and are described in the following CWC article: Recent Modifications of Imputation Methods for National Compensation Survey Benefits Data, found at the following link: <http://www.bls.gov/opub/cwc/cm20090825ar01p1.htm>.

**3c. Non-Response Bias Research**

Recently, 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 referenced in Section 2c, by Ponikowski, McNulty, and Crockett, and listed in the 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. Implementation of the new pages will follow successful testing by the BLS Cognitive Laboratory, both in the laboratory and then in the field. Respondent access to these new web pages will be through our existing IDCF system.

Through environmental scanning, NCS has identified Payroll Deduction IRA plans as an upcoming compensation trend. NCS plans to add this as a subcategory of the currently collected Cash or Deferred Arrangements (CODAs) with no employer contributions and measure access to the new benefit. This collection will contain yes/no/not determinable questions on whether an establishment offers the new benefit. Our research has shown that staff understand the new benefit and its relationship to the existing benefit-- Cash or Deferred Arrangements (CODAs) with no employer contributions.

A nonsubstantive change will be submitted to OMB for both the Web pages and the new subcategory of the currently collected cash or Deferred Arrangements (CODAs) when testing is complete.

**4b. Tests of Survey Design Procedures**

As mentioned previously, the President’s proposed budget for FY 2011 calls for an alternative to the LPS component of the NCS, a new approach that uses data from two current BLS programs – the 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. If this change is approved for implementation, the NCS will need to be redesigned. Planning for a redesigned NCS is underway. These plans, currently being developed, 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 will implement a model-based estimation approach to produce data for the President’s Pay Agent. NCS may also 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 is currently evaluating and testing alternatives for this change in three separate sets of activities.

First, the BLS staff is examining 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 is testing 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 are also studying the effect of any redesign on response rates and bias.

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.

Results of these tests will be used to determine which changes will be made to the sample design. Until the proposed budget changes are implemented, testing is complete, specific changes are identified, and the modified design is approved by OMB, NCS will continue to use the sample design and rotation strategy described earlier in this document.

Second, the BLS staff explored and evaluated different model-based approaches that use data from the OES survey and the Employment Cost Index 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 titled “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 C).

As described in Attachment C, the BLS would introduce a new model-based approach that uses data from two current BLS programs – the Occupational Employment Statistics (OES) survey and ECI, another component of the NCS. In the proposed approach, OES data would provide wage data by occupation and by area, while NCS data would be used to specify grade level effects. Since the OES sample is much larger than the NCS sample, the BLS would expect, overall, efficiency gains in the estimates of mean wages by occupation and area. The model-based approach also could be used to extend the estimation of pay gaps to areas that are not present in the NCS sample. This approach also would allow the BLS to eliminate 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 C, 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.

Third, BLS staff is 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 C).

Gwyn Ferguson, Chester Ponikowski, Joan Coleman, (August 2010), “Evaluating Sample Design Issues in the National Compensation Survey,” ASA Papers and Proceedings, (See Attachment D).

**Appendix A: Allocation of NCS Establishment Sample by Survey Area and Ownership, followed by Private Industry In-Scope Emp & Estabs by Survey Area (see below). Please note, total allocated sample shown in table does not include 50 establishments sampled independently from the aerospace industry.**

****

****

| In-scope employment | NCS Area | In-scope establishments |
| --- | --- | --- |
| 2,048,222 | Atlanta-Sandy Springs-Gainesville, GA-AL CSA | 142,984 |
| 2,622,436 | Boston-Worcester-Manchester, MA-NH CSA | 161,991 |
| 463,840 | Buffalo-Niagara-Cattaraugus, NY CSA | 28,560 |
| 3,861,672 | Chicago-Naperville-Michigan City, IL-IN-WI CSA | 244,637 |
| 893,908 | Cincinnati-Middletown-Wilmington, OH-KY-IN CSA | 49,072 |
| 1,181,636 | Cleveland-Akron-Elyria, OH CSA | 74,253 |
| 824,733 | Columbus-Marion-Chillicothe, OH CSA | 44,933 |
| 2,565,754 | Dallas-Fort Worth, TX CSA | 142,314 |
| 380,299 | Dayton-Springfield-Greenville, OH CSA | 22,268 |
| 1,193,645 | Denver-Aurora-Boulder, CO CSA | 95,483 |
| 1,836,030 | Detroit-Warren-Flint, MI CSA | 111,840 |
| 545,693 | Hartford-West Hartford-Willimantic, CT CSA | 33,151 |
| 2,232,138 | Houston-Baytown-Huntsville, TX CSA | 119,698 |
| 205,012 | Huntsville-Decatur, AL CSA | 12,680 |
| 839,031 | Indianapolis-Anderson-Columbus, IN CSA | 48,436 |
| 5,901,695 | Los Angeles-Long Beach-Riverside, CA CSA | 390,188 |
| 550,909 | Louisville-Elizabethtown-Scottsburg, KY-IN CSA | 33,699 |
| 794,265 | Milwaukee-Racine-Waukesha, WI CSA | 42,222 |
| 1,647,299 | Minneapolis-St. Paul-St. Cloud, MN-WI CSA | 100,571 |
| 8,168,608 | New York-Newark-Bridgeport, NY-NJ-CT-PA CSA | 615,164 |
| 2,370,622 | Philadelphia-Camden-Vineland, PA-NJ-DE-MD CSA | 155,847 |
| 1,001,297 | Pittsburgh-New Castle, PA CSA | 62,094 |
| 648,791 | Raleigh-Durham-Cary, NC CSA | 41,047 |
| 668,957 | Sacramento--Arden-Arcade--Truckee, CA-NV CSA | 47,057 |
| 2,870,852 | San Jose-San Francisco-Oakland, CA CSA | 181,538 |
| 1,491,433 | Seattle-Tacoma-Olympia, WA CSA | 106,947 |
| 3,338,032 | Washington-Baltimore-No. Virginia, DC-MD-VA-WV CSA | 224,372 |
| 326,082 | Albany-Schenectady-Troy, NY | 21,267 |
| 362,495 | Albuquerque, NM | 20,276 |
| 287,025 | Allentown-Bethlehem-Easton, PA-NJ | 18,794 |
| 396,157 | Amarillo, TX | 5,591 |
| 409,285 | Atlantic City, NJ | 6,603 |
| 381,426 | Auburn-Opelika, AL | 2,457 |
| 605,334 | Austin-Round Rock, TX | 38,251 |
| 428,106 | Bangor, ME | 3,913 |
| 294,632 | Baton Rouge, LA | 19,805 |
| 598,277 | Billings, MT | 6,007 |
| 409,746 | Birmingham-Hoover, AL | 26,823 |
| 282,495 | Bloomington, IN | 3,683 |
| 418,468 | Bloomington-Normal, IL | 3,413 |
| 374,851 | Brainerd, MN | 2,922 |
| 438,225 | Brownsville-Harlingen, TX | 5,871 |
| 490,680 | Cedar Rapids, IA | 7,081 |
| 524,374 | Centralia, WA | 1,735 |
| 454,636 | Charleston-North Charleston, SC | 16,281 |
| 720,529 | Charlotte-Gastonia-Concord, NC-SC | 45,064 |
| 231,076 | Claremont, NH | 1,062 |
| 384,420 | Clarksburg, WV | 2,109 |
| 411,878 | Columbia, SC | 16,227 |
| 339,107 | Corning, NY | 1,849 |
| 463,960 | Corpus Christi, TX | 9,001 |
| 306,494 | Elkhart-Goshen, IN | 4,841 |
| 409,348 | El Paso, TX | 12,780 |
| 318,240 | Emporia, KS | 899 |
| 421,222 | Fayetteville, NC | 5,982 |
| 336,155 | Fort Collins-Loveland, CO | 9,999 |
| 372,479 | Fort Walton Beach-Crestview-Destin, FL | 5,936 |
| 384,060 | Fresno, CA | 14,803 |
| 327,417 | Grand Rapids-Wyoming, MI | 16,177 |
| 412,880 | Great Falls, MT | 2,579 |
| 427,781 | Greensboro-High Point, NC | 17,312 |
| 425,201 | Greenville, SC | 15,069 |
| 358,326 | Hickory-Lenoir-Morganton, NC | 7,442 |
| 331,782 | Holland-Grand Haven, MI | 5,403 |
| 351,868 | Honolulu, HI | 23,912 |
| 409,015 | Iowa City, IA | 3,924 |
| 430,345 | Jackson, MS | 12,685 |
| 514,638 | Jacksonville, FL | 38,642 |
| 405,874 | Johnstown, PA | 3,468 |
| 416,309 | Kalispell, MT | 4,266 |
| 834,907 | Kansas City, MO-KS | 53,824 |
| 395,132 | Kennewick-Richland-Pasco, WA | 4,895 |
| 489,305 | Knoxville, TN | 15,576 |
| 387,956 | Lafayette, LA | 9,401 |
| 260,970 | Lancaster, SC | 1,053 |
| 779,190 | Las Vegas-Paradise, NV | 48,971 |
| 462,116 | Lincoln, NE | 8,163 |
| 404,121 | Little Rock-North Little Rock, AR | 20,592 |
| 343,828 | Logansport, IN | 690 |
| 408,325 | Madison, WI | 14,567 |
| 440,651 | Manitowoc, WI | 1,688 |
| 385,980 | Meadville, PA | 2,099 |
| 396,539 | Medford, OR | 6,149 |
| 516,686 | Memphis, TN-MS-AR | 24,606 |
| 550,011 | Miami, OK | 575 |
| 1,955,368 | Miami-Fort Lauderdale-Miami Beach, FL | 195,749 |
| 348,922 | Mobile, AL | 9,369 |
| 391,202 | Monroe, LA | 4,752 |
| 321,058 | Mount Airy, NC | 1,628 |
| 331,444 | Murray, KY | 841 |
| 368,967 | Muskegon-Norton Shores, MI | 3,011 |
| 424,653 | Muskogee, OK | 1,406 |
| 628,861 | Nashville-Davidson--Murfreesboro, TN | 35,982 |
| 438,541 | New Orleans-Metairie-Kenner, LA | 34,698 |
| 380,529 | Nogales, AZ | 1,184 |
| 392,962 | Ocala, FL | 8,105 |
| 450,069 | Oklahoma City, OK | 32,496 |
| 393,617 | Omaha-Council Bluffs, NE-IA | 22,993 |
| 891,345 | Orlando, FL | 62,181 |
| 413,537 | Ottumwa, IA | 829 |
| 375,740 | Paducah, KY-IL | 2,702 |
| 340,081 | Palatka, FL | 1,426 |
| 469,358 | Palm Bay-Melbourne-Titusville, FL | 14,603 |
| 1,556,862 | Phoenix-Mesa-Scottsdale, AZ | 100,979 |
| 856,908 | Portland-Vancouver-Beaverton, OR-WA | 66,747 |
| 581,962 | Providence-New Bedford-Fall River, RI-MA | 44,387 |
| 442,489 | Quincy, IL-MO | 1,900 |
| 409,423 | Reading, PA | 8,683 |
| 394,916 | Reno-Sparks, NV | 14,244 |
| 491,711 | Richmond, VA | 33,755 |
| 452,140 | Roanoke, VA | 8,303 |
| 414,382 | Rochester, NY | 23,493 |
| 425,676 | Rockford, IL | 7,499 |
| 1,112,832 | St. Louis, MO-IL | 67,693 |
| 370,035 | Salem, OR | 9,507 |
| 400,786 | Salinas, CA | 7,979 |
| 432,819 | Salisbury, MD | 3,041 |
| 540,816 | Salt Lake City, UT | 40,323 |
| 678,406 | San Antonio, TX | 37,748 |
| 1,061,305 | San Diego-Carlsbad-San Marcos, CA | 71,286 |
| 363,986 | Sarasota-Bradenton-Venice, FL | 23,822 |
| 406,312 | Sioux City, IA-NE-SD | 3,938 |
| 412,066 | Springfield, MA | 14,697 |
| 465,067 | Springfield, MO | 10,668 |
| 294,624 | Starkville, MS | 787 |
| 290,142 | State College, PA | 3,201 |
| 299,235 | Tallahassee, FL | 9,195 |
| 1,004,489 | Tampa-St. Petersburg-Clearwater, FL | 81,663 |
| 351,719 | Toledo, OH | 14,890 |
| 417,485 | Tucson, AZ | 19,654 |
| 366,340 | Tulsa, OK | 23,734 |
| 342,777 | Tuscaloosa, AL | 4,417 |
| 578,039 | Virginia Beach-Norfolk-Newport News, VA-NC | 38,436 |
| 323,420 | Visalia-Porterville, CA | 5,716 |
| 379,831 | Wausau, WI | 3,133 |
| 457,490 | Wilmington, NC | 9,893 |
| 390,468 | Wooster, OH | 2,425 |
| 472,175 | York-Hanover, PA | 8,599 |
| 421,110 | Youngstown-Warren-Boardman, OH-PA | 13,337 |
| 424,208 | Atchison, Holt Counties MO & Johnson, Nemaha, Otoe, Pawnee, Richardson Counties NE (Southeastern Nebraska-Northwestern Missouri) | 1,403 |
| 245,730 | Baylor, Briscoe, Childress, Cottle, Dickens, Floyd, Foard, Hall, Hardeman, Haskell, Kent, King, Knox, Motley, Stonewall, Throckmorton Counties, TX (Northwest Texas) | 1,145 |
| 203,336 | Bedford, Fulton, Juniata Counties PA | 1,797 |
| 203,492 | Caledonia and Orleans Counties, VT | 1,733 |
| 393,168 | Carroll and Jo Daviess Counties, IL & Lafayette County, WI | 1,450 |
| 458,821 | Claiborne, Franklin, Jefferson and Wilkinson Counties, MS (Southwestern Mississippi) | 393 |
| 437,908 | Esmeralda, Lyon and Mineral Counties, NV | 1,100 |
| 444,309 | Fannin, Gilmer, and Lumpkin Counties, GA | 1,628 |
| 316,406 | Fayette, Lee Counties, TX | 1,051 |
| 151,459 | Ferry, Okanogan Counties, WA | 1,248 |
| 382,662 | Graham, Norton, Osborne, Phillips, Rooks, Smith Counties, KS (North Central Kansas) | 874 |
| 435,842 | Lee, Norton City, Wise Counties, VA | 1,606 |
| 342,918 | Sanilac County, MI | 842 |
| *108,011,978* | *Total – All Survey Areas* | *5,328,096* |

**Appendix B: NCS Sample Rotation**

See next page.



Attachment C



Attachment D



1. Yoel Izsak, Lawrence R. Ernst, Erin McNulty, Steven P. Paben, Chester H. Ponikowski, Glenn Springer, Jason Tehonica, "Update on the Redesign of the National Compensation Survey", 2005 Proceedings of the American Statistical Association, Section on Survey Methods Research [CD-ROM], American Statistical Association, 2005 <http://www.bls.gov/ore/pdf/st050140.pdf>> [↑](#footnote-ref-1)
2. 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> [↑](#footnote-ref-2)