Employment, Wages, and Contributions Report (QCEW Program)

B. COLLECTION OF DATA EMPLOYING STATISTICAL METHODS

1a. Universe

The universe of respondents to U.S. Bureau of Labor Statistics for Quarterly Census of Employment and Wages are the 50 States, the District of Columbia, Puerto Rico, and the U.S. Virgin Islands. The ultimate source of data for these 53 entities is the Quarterly Contribution Reports (QCR) submitted to State Workforce Agencies (SWAs) by employers subject to State Unemployment Insurance (UI) laws. The QCEW data, which are compiled for each calendar quarter, provide a comprehensive business name and address file with employment and wage information by industry, at the six-digit North American Industry Classification System (NAICS) level, and at the national, State, Metropolitan Statistical Area, and county levels for employers subject to State UI laws. Similar data for Federal Government employees covered by the Unemployment Compensation for Federal Employees program (UCFE) also are included.

The QCEW program provides a virtual census of nonagricultural employees and their wages, with about 51 percent of the workers in agriculture covered as well. As shown in table 1, the number of private business establishments (worksites) covered is about 8.66 million and the covered employment is about 115 million. Additionally, about 64,000 Federal government, 67,000 State government, and 160,000 Local government establishments are covered; private households and all agriculture except forestry are excluded. The total number of covered establishments is about 8.95 million and the covered employment is about 137 million in June 2007. The QCEW series has broad economic significance in measuring labor trends and major industry developments, in time series analyses and industry comparisons, and in special studies such as analyses of establishments, employment, and wages by size of establishment.

The BLS role in the QCEW program is to establish and enforce uniform methods and processes that yield consistent level of data quality for the multifaceted uses of the QCEW data. The BLS role is to take in raw U.I. administrative data, to understand error components, to address each with methods and processes to reduce resulting error and yield high quality economic data and sample frame. The improvement processes include but are not limited to: efficiency in data collection from large multi-establishment employers through Electronic Data Interchange (EDI); statistically valid procedures for editing, estimating missing reports and data elements, and record linkage; and standardized processing systems, training of staff; and quality controls procedures for data review (see sections 2b and 2c on estimation procedures and reliability for details). After the data have gone through extensive review at the state, regional, and national levels, the BLS summarizes these data to produce totals for all counties, Statistical Areas, the States, and the Nation by various industrial levels.

1b. Sample

QCEW is a census of establishments; hence, every unit is in the sample and represents itself only. That is, each unit has a sampling weight of 1.00.

2a. Sample Design

The QCEW is a census. The sample design for QCEW is very simple since all establishments are included with a sampling weight of 1.00 or with certainty. The sampling unit is the establishment or worksite.

2b. Estimation Procedure

The aggregated totals of employment and wages for each sub-domain (e.g., industry, geography, and size) are simply the sum of the micro records belonging to that sub-domain. Averages and other statistics for each sub-domain are derived by performing the appropriate arithmetic functions.

As mentioned above, the BLS role is to add quality to the raw data. One of these processes involves editing the data and conducting validation checks. The basic monthly employment edit consists of a six step statistical test that includes the use of multiple t-test for: month-to-month change, over-the-year change, and a 12 month variation in data; some tests are conducted on levels while others are conducted on rate of change. The wage edit includes the use of inter-quartile test developed by Hoaglin, Iglewicz, and Tukey. The Edit Conditions and Formulas are described in Appendix-F of the QCEW Operating Manual (2007).

Although BLS receives QCEW files from all 53 entities in a timely manner, the files contain estimates for late and missing respondents. Therefore, a step in the data process is estimation for late respondents and for missing respondents (i.e., unit non-response) and data elements (i.e., item non-response). As shown in table 2a, about five percent of the establishments respond late or fail to respond to QCEW in a timely manner; the corresponding figure for employment is about three percent as shown in table 2b. The non-response rates for wages are about four percent as shown in table 2c.

The current method of imputation applies the missing establishment a-year-ago change to the previous month employment or quarterly wages to estimate the current month employment or quarterly wages. That is, missing establishment current month's employment is equal to the previous month employment multiplied by its a-year-ago change; similar procedure is applied to estimate total quarterly wages. A drawback to this procedure is that it uses a-year-ago trend rather than current trend. The current Imputation Formulas are described in Chapter 8 and Appendix-J of the QCEW Operating Manual.

BLS is currently conducting extensive research on alternative imputation methods for both employment and wages. The alternative methods use current trends. They are: 1) a ratio method which uses rate of change from previous month (quarter) to the current month (quarter); and 2) the "hot-deck nearest neighbor" method which uses respondents in similar cells as non-respondents to fill-in or as donors for the missing data. This report should be available in late 2008.

Another data processing step is to link the QCEW data across quarters for various purposes including: 1) editing and imputation; 2) separation of establishments into new establishments (openings or births), continuous establishments (existing businesses), and

out-of-business establishments (closings or deaths); and 3) longitudinal research. The BLS has employed the Felligi and Sunter record linkage methodology. See the paper by Kenneth Robertson, et al. (1997).

2c. Reliability

Since QCEW is a census, the data are only subject to non-sampling errors. To control for these non-sampling errors, BLS has put in place extensive quality control procedures that include: 1) improved data collection methods especially for large multi-establishment employers through Electronic Data Interchange; 2) standardized data processing systems that include edits, imputation, record linkages and industrial classification coding; and 3) standardized training of staff at state, regional, and national levels in the review of data according to the guidelines provided by the QCEW policy council and stated in official memorandums (available upon request). Records that fail these edits are individually reviewed. Respondent contact is frequently used to validate significant movements or to correct the data.

The two most important initiatives undertaken by BLS to enhance the quality of QCEW data are the establishment of the Multiple Worksites Report (MWR) Survey and the Annual Re-filing Survey (ARS). Two separate OMB clearances are obtained for these Surveys. The MWR form is sent quarterly to multi-establishment employers for the purpose of asking them to break out their consolidated reports to the establishment level. For example, some employers provide data for all of their operations within a State or at the county level; the MWR asks the employer to provide information for each establishment so that all records on the file can be at the establishment level which is generally the sampling unit for most of the BLS surveys. This also improves the quality of local economic data by more accurately reporting the location and type of economic activity.

The Annual Re-filing Survey is conducted annually on about one-fourth of the establishments on the frame for the purpose of updating the industrial classification, business name, reporting and physical location addresses, and auxiliary status. These establishments are selected randomly. The states and regional staff are trained extensively in the industrial classification coding. Additionally, standardized systems are provided to the states and regions to process the data.

2d. Revisions

For the first quarter of each year, QCEW data are published five times; the original data are first released in October of the same year followed by revisions in January, April, July, and October of the following year. For example, March 2006 data are first published in October 2006, then in January, April, July, and October of 2007. The 2nd quarter data is published four times; the 3rd quarter data is published three times; and the 4th quarter data is published twice. Table 3a provides data for the initial publication of each quarter in 2006 to their final publication in October 2007. As shown in table 3b, the largest revision generally occurs from initial publication to the first revision as missing reports for late responding employers come in including out-of-business reports. The magnitude of revisions is relatively small; that is, less than 0.1 percentage point.

2e. Specialized Procedures

None

2f. Data Collection Cycles

The QCEW program is quarterly as the employers are required to file Quarterly Contribution reports (UI reports) on a quarterly basis.

3. Methods to Maximize Response Rates

Since employers are required to file Quarterly Contributions Reports under the UI law for each state, the response rates are generally very high. The unit response rates for employment are about 95 percent (table 2a) and about 96 percent (table 2c) for wages as reporting of wages are required by UI law. The response rates based on total covered employment are about 97 percent (table 2b) as the non-response is mostly concentrated among the small establishments.

Growth of EDI, the direct transfer of data from the firm to BLS, also provides a high level of response and stability. BLS currently collects over 80,000 reports from nearly 100 large firms with about 10 million employees via EDI. For final estimates, virtually all of these firms provide data.

4. Tests

BLS has undertaken several new initiatives in the area of research on control and measurement of non-sampling error. The 1991 benchmark of Current Employment Statistics Survey's estimate of employment to the QCEW revealed a substantial non-sampling error problem caused by payroll processing firms. The American Statistical Association formed a committee to review BLS procedures and issued a report in January 1994 (American Statistical Association, 1994). BLS has adopted most of the report's recommendations. BLS has also conducted a Response Analysis Survey of Payroll Processing Firms (Goldenberg, Moore, and Rosen, 1994). The purpose of the survey was to identify practices that can affect the data collected by the CES program and the Quarterly Census of Employment and Wages program (the benchmark source data) and educate payroll processors on proper reporting procedures. BLS has also conducted a Response Analysis Survey (RAS) of CES and QCEW covered employment reporting (Werking, Clayton, and Rosen, 1995). The survey identified factors that affect both CES and QCEW reporting within the same firm. Based on these RAS studies, BLS has undertaken an extensive education program with CES respondents. This includes highlighting correct reporting of problem items on the CES report form and the inclusion of special notices on correct reporting on the monthly advance notice fax message. A new RAS is planned for collection in 2008 and a report detailing new findings should be available in late 2008.

5. Statistical and Analytical Responsibility

Ms. Shail Butani, Chief, Statistical Methods Division of the Office of Employment and Unemployment Statistics, is responsible for the statistical aspects of the QCEW program. Ms. Butani can be reached on 202-691-6347. As mentioned in the above paragraph, BLS seeks consultation with other outside experts on an as needed basis. The QCEW Policy council has been consulted on the content, uses, and methodology of the program. The Policy Council is composed of ten representatives of SWAs and BLS staff.

6. References

American Statistical Association (1994) "A Research Agenda to Guide and Improve the Current Employment Statistics Survey." American Statistical Association Panel for the Bureau of Labor Statistics' Current Employment Statistics Survey, January, 1994. Alexandria, VA: American Statistical Association (available upon request).

Bureau of Labor Statistics. BLS Handbook of Methods Chapter 5: Employment and Wages Covered by Unemployment Insurance. Washington DC: Bureau of Labor Statistics, 2004, p.42-47. http://www.bls.gov/opub/hom/pdf/homch5.pdf http://www.bls.gov/opub/hom/homch5_d.htm

Bureau of Labor Statistics. Official memorandums to the States and Regional staff on QCEW program. (available upon request).

David C. Hoaglin, Boris Iglewicz, John W Tukey (1996) "Performance of Some Resistant Rules for Outlier Labeling." Journal of the American Statistical Association, Vol. 81 No. 396. (Dec., 1986), pp 991-999.
 http://www.jstor.org/view/01621459/di985970/98p05363/0

Edit Conditions and Formulas. Appendix-F QCEW Operating Manual(2007). Bureau of Labor Statistics, Washington, DC-20212.(available on CD).

Fellegi, I. P. and Sunter, A. B. (1969). A theory for record Linkage, Journal of the American Statistical Association, 64, 1183-1210.

Goldenberg, Karen L., Susan E. Moore, and Richard J. Rosen (1994) "Commercial Payroll Software and the Quality of Employment Data." Proceedings of the Survey Research Methods Section, American Statistical Association, 13-18 August, 1994. Toronto: American Statistical Association, 1994.
http://www.amstat.org/sections/SRMS/Proceedings/papers/1994_178.pdf

Imputation Formulas. Chapter 8 and Appendix J, QCEW Operating Manual(2007). Bureau of Labor Statistics, Washington, DC-20212. (available on CD).

Kenneth Robertson, Larry Huff, Gordon Mikkelson, Timothy Pivetz, and Alice Winkler (1997). "Improvement in Record Linkage Processes for the Bureau of Labor Statistics' Business Establishment List." In Record Linkage Techniques (1997). Proceedings of an International Workshop and Exposition. Edited by; Wendy Alvey and Bettye Jamerson, Federal Committee on Statistical Methodology, Office of Management and Budget, Washington, DC.

Werking, George S., Richard L. Clayton, and Richard J. Rosen (1995) "Studying the Causes of Employment Count Differences Reported in Two BLS Programs." Proceedings of the Survey Research Methods Section, American Statistical Association, 13-17 August, 1995. Orlando: American Statistical Association, 1995.
http://www.amstat.org/sections/SRMS/Proceedings/papers/1995_137.pdf

Table 1--QCEW Survey Universe for 50 States, D.C., Puerto Rico, and Virgin Island on NAICS basis				
(April, May, June 2007)(in thousands)				
Description	No. of Establishments	April 2007	May 2007	June 2007
Total	**8,946**	**134,917**	**136,200**	**137,018**
Total Private	**8,655**	**113,175**	**114,342**	**115,503**
Agriculture, forestry, fishing and hunting	95	1,143	1,213	1,287
Mining	29	648	656	668
Utilities	16	547	548	554
Construction	889	7,447	7,654	7,835
Manufacturing	361	13,827	13,852	13,954
Wholesale trade	624	5,953	5,982	6,024
Retail Trade	1,048	15,248	15,413	15,496
Transportation and Warehousing	221	4,248	4,292	4,314
Information	143	3,009	3,032	3,055
Finance and insurance	491	6,005	6,005	6,025
Real estate and rental and leasing	376	2,135	2,157	2,193
Professional, Scientific and Technical Services	965	7,632	7,536	7,615
Management of companies and enterprises	48	1,826	1,832	1,849
Administrative and support and waste management services	455	8,298	8,476	8,564
Educational services	85	2,340	2,328	2,209
Health care and social assistance	732	15,012	15,084	15,167
Arts, entertainment, and recreation	124	1,903	2,035	2,201
Accommodation and food services	598	11,299	11,522	11,687
Other services, except public administration	1,138	4,408	4,454	4,517
Unclassified	215	248	271	290
Federal Government	**64**	**2,718**	**2,727**	**2,743**
State Government	**67**	**4,643**	**4,636**	**4,577**
Local Government	**160**	**14,382**	**14,494**	**14,196**

	Table 2a. U.S. Percentage of imputed establishments by year and month											
	Jan	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
2001	5.59	5.61	5.64	5.41	5.44	5.50	4.61	4.65	4.67	4.60	4.65	4.69
2002	5.14	5.18	5.17	4.71	4.73	4.79	4.61	4.63	4.68	4.36	4.42	4.46
2003	5.88	5.93	5.92	5.21	5.20	5.26	4.81	4.82	4.84	5.04	5.08	5.13
2004	5.51	5.54	5.52	5.38	5.38	5.47	5.00	5.01	5.11	4.79	4.83	4.91
2005	5.09	5.12	5.16	4.66	4.65	4.79	4.81	4.83	4.83	4.17	4.25	4.32
2006	5.48	5.52	5.53	4.50	4.47	4.54	4.31	4.41	4.45	4.50	4.67	4.72

	Table 2b. U.S. Percentage of imputed employment by year and month											
	Jan	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
2001	3.18	3.18	3.18	3.16	3.17	3.16	2.79	2.81	2.84	3.27	3.27	3.31
2002	2.87	2.87	2.85	2.89	2.87	2.86	3.00	3.00	3.02	2.97	2.96	2.95
2003	3.43	3.46	3.36	2.90	2.86	2.84	3.12	3.11	3.12	3.13	3.12	3.14
2004	2.93	2.89	2.85	3.15	3.08	3.12	2.91	2.87	3.02	2.81	2.78	2.83
2005	2.49	2.50	2.51	2.46	2.42	2.54	2.51	2.52	2.48	2.36	2.36	2.37
2006	2.40	2.42	2.39	1.89	1.86	1.87	1.93	1.96	1.97	2.32	2.39	2.38

Table 2c: Units imputed for Missing Wages			
Year/qtr	Imputed units	Total Units	%of imputed units
2006-1st	410,596	8,752,182	4.69
2006-2nd	340,064	8,757,775	3.88
2006-3rd	317,630	8,824,858	3.60
2006-4th	275,654	8,913,853	3.09
2007-1st	373,835	8,927,738	4.19
2007-2nd	394,328	8,927,961	4.42

Note: The data analysis does not include Puerto Rico and US Virgin islands.

Table 3a: Revisions in published data, U.S. total

March 2006 (October 2006 release)	March 2006 (January 2007 Release)	March 2006 (April 2007 Release)	March 2006 (July 2007 Release)	March 2006 (October 2007 Release)	First revision	Second revision	Third revision	Fourth revision	Total Revision (since October 2006)
132,613,061	132,557,606	132,538,199	132,518,588	132,518,615	-55,455	-19,407	-19,611	27	-94,446

	June 2006 (January 2007 Release)	June 2006 (April 2007 Release)	June 2006 (July 2007 Release)	June 2006 (October 2007 Release)	First revision	Second revision	Third revision		Revision since January '2007
	135,481,097	135,441,269	135,424,126	135,419,165	-39,828	-17,143	-4,961		-61,932

		September 2006 (April 2007 Release)	September 2006 (July 2007 Release)	September 2006 (October 2007 Release)	First revision	Second revision			Revision since April '2007
		134,988,878	135,025,324	135,030,283	36,446	4,959			41,405

			December 2006 (July 2007 Release)	December 2006 (Oct-2007 Release)	First revision				Revision
			135,933,219	135,938,194	4,975				

Table 3b: Percentage of revision from original publication				
Preliminary Publication	March 2006 (January 2007 release)	June 2006 (April 2007 release)	September 2006 (July 2007 release)	December 2006 (October 2007 release)
%Revision from preliminary published data	-0.0712	-0.0457	0.0307	0.0037