THE RELIABILITY OF THE LOCALITY PAY ESTIMATES

Joan L. Coleman & Penny L. James, Bureau of Labor Statistics
Penny L. James • PSB Room 3160 • 2 Mass. Ave., NE • Washington, DC 20212

INTRODUCTION

In 1990 Congress passed the Federal Employees Pay Comparability Act (FEPCA). One of the purposes of this law was to change the way in which federal salaries were adjusted, from a national to locality-based adjustment. The Bureau of Labor Statistics (BLS), in support of the FEPCA, was required to develop and maintain a program of collection, compilation, and analysis of occupational wages in private industry and state and local governments. The program was to be designed to yield national estimates for selected occupations, and comparable estimates for selected metropolitan areas.

Current programs within the BLS were not designed to supply the necessary wage information that the Office of Personnel Management (OPM) required for the FEPCA. To accommodate OPM's needs, the BLS integrated and expanded two already existing programs, the White Collar Pay and Area Wage Survey programs. By integrating these two programs, the BLS was able to obtain the information necessary for the FEPCA, while maintaining and improving its existing programs.

White Collar Pay Survey Program

The White Collar Pay Program was designed to provide national estimates of wage information for 147 work level categories in 28 different occupations. These occupational work levels were defined to include only those workers meeting specific criteria as to training, job functions, and responsibilities. The types of occupations surveyed were classified as (1) professional and administrative; (2) technical support; and (3) clerical. This survey covered establishments in the United States, except Hawaii and Alaska, employing at least fifty workers in the goods and service producing industries. The survey was conducted on a two-year cycle with service producing industries surveyed in odd number years and goods producing in even number years. The service producing industries included Transportation, Communications, and other Public Utilities; Wholesale Trade; Retail Trade; Finance, Insurance and Real Estate; and Services. The goods producing industries included Manufacturing, Construction and Mining.

Area Wage Survey Program

Before the integration, the Area Wage Survey program was a locality based program that provided wage and related benefit information for representative establishments within six broad industry divisions: Manufacturing; Transportation, Communication and other Public Utilities; Wholesale Trade; Retail Trade; Finance, Insurance and Real Estate; and Services. The program consisted of ninety areas, which were selected to represent the 326 metropolitan statistical areas in the contiguous U.S. Thirty-two of these, which were selected with certainty based on non-agricultural employment, were surveyed annually. Two groups of 29 areas were surveyed in alternate years and were chosen to represent the remainder of the metropolitan areas. Occupations selected for study were common to a variety of manufacturing and nonmanufacturing industries, and were of the following types: (1) office clerical; (2) professional and technical; (3) maintenance, toolroom, and powerplant; and (4) material movement and custodial.

Integration

In order to meet the needs of the FEPCA, the OPM requested information on the 28 Consolidated Metropolitan Statistical Areas (CMSAs) that contain the largest proportion of civilian federal employees. The OPM also wanted estimates to represent the rest of the United States. To provide this information, the BLS decided to combine aspects of both the WCP and the AWS programs. The sampling design, which was cluster based, was taken from the AWS program. Forty-six of the ninety AWS areas were part of the 28 OPM pay districts (CMSAs). Thirty-two of these areas, which were called critical areas, were surveyed to represent the CMSAs. Because the other fourteen of these 46 could not represent the rest of the U.S., other areas were chosen as replacements to insure the statistical purity of the rest of the U.S. estimates. These statistical purity areas along with the remaining 42 of the ninety AWS areas represented the metropolitan part of the rest of the U.S. To represent the nonmetropolitan area, seventy out of 2383 nonmetropolitan counties were surveyed also.

This new combined program is referred to as the Occupational Compensation Survey Program (OCSP) or locality pay. It uses the WCP industrial scope which included more service producing and goods producing Standard Industrial Classifications (SICs) than the AWS program. The combined program also expands the scope to include state and local governments, which was not previously included in either the WCP or AWS programs. As with both the WCP and the AWS programs, the minimum establishment employment for the OCSP is fifty employees.

The job list for the OCSP is also a combination of the two programs, but does not contain all jobs in both programs. Emphasis is placed on professional and administrative occupations. Other types of occupations studied are technical; protective service; clerical; maintenance and toolroom; and material movement and custodial.

The OCSP was developed not only to satisfy the requirements of the OPM for locality pay, but also to maintain as much as possible the BLS survey programs. In so doing, not only is the BLS able to satisfy the needs of the OPM, but also through expanded scope and occupations, improve its WCP national and AWS local publications.

RELIABILITY

The plan in integrating the WCP and AWS programs into the locality pay was to create a program that would be an improvement over each individual program through the expanded scope. However, the effect this would have on the reliability of the estimates was never measured in actual surveys.

Basically three sets of estimates were produced in the locality pay program. First, there were 32 primary metropolitan statistical areas that represent the 28 pay districts for which the OPM requested local estimates. These 32 areas are referred to as critical areas. Second, there were several other metropolitan and nonmetropolitan areas that were combined to produce estimates on the rest of the United States. Finally, national estimates were produced by combining the 32 critical areas with the rest of the U.S.

To determine the reliability of the OCSP, relative standard errors (RSEs) for the mean wage estimates for all critical areas, the rest of the U.S., and the national product were compared to previously calculated RSEs for surveys in the WCP and AWS programs. For all of the OCSP and AWS surveys, the RSE calculations were based on the closed form variance estimation procedure for estimating totals and means over subpopulations. For WCP, the RSEs were calculated using the random group method of variance estimation. The WCP RSEs were calculated using a replication technique with fifteen random groups.

National OCSP versus White Collar Pay

Below are the RSE distributions at the 68 percent level in private industry for the OCSP national and WCP surveys. The distributions for the WCP survey are based on the 1990 bulletin, which combined the 1989 service producing data, updated by an Employment Cost Index factor, with the 1990 goods producing data. The distributions represent all publishable RSEs in each of the surveys.

	National		White Collar Pay	
	Dist.	Perc.	Dist.	Perc.
Less than 1%	41	35.0	72	52.6
1 and under 2%	55	47.0	51	37.2
2 and under 3%	17	14.5	9	6.6
3% or more	4	3.5	5	3.6
	117		137	

The table above shows that the distributions of RSEs for both surveys are very similar. The one obvious difference is that a larger percentage of RSEs in the WCP survey fell in the less than 1% category. This disparity is partly explained by the differences in occupational coverage. The WCP survey published more occupational levels, 137 compared to 117 in the national OCSP. All of these additional jobs were white collar occupations, which tend to have smaller RSEs.

Also, the OCSP produced national estimates for the first time for blue collar occupations. Twenty-two occupational levels in the locality pay program's national survey were classified as maintenance, toolroom, material movement and custodial. The blue collar occupations typically have larger RSEs than the white collar jobs. The larger RSEs are due to bimodal distributions caused by large differences in wages for occupations that could be either union or nonunion depending on the establishment. The table below eliminates these 22 occupations from the OCSP national distributions. As can be seen, these distributions are now more similar between the OCSP national and WCP RSEs.

	National		White Collar Pay	
	Dist.	Perc.	Dist.	Perc.
Less than 1%	41	43.2	72	52.6
1 and under 2%	43	45.3	51	37.2
2 and under 3%	9	9.5	9	6.6
3% or more	2	2.1	5	3.6
	95		137	

Rest of the U.S. versus White Collar Pay

The table below shows the comparison of the RSE distributions for average earnings at the 68 percent level between the rest of the U.S. and White Collar Pay. The RSEs for WCP are again smaller than those from the rest of the U.S.

	Rest of U.S.		White Collar Pay	
	Dist.	Perc.	Dist.	Perc.
Less than 1%	16	17.6	72	52.6
1 and under 2%	40	44.0	51	37.2
2 and under 3%	24	26.4	9	6.6
3% or more	11	12.1	5	3.6
	91		137	

The table below compares the WCP RSEs to the OCSP rest of the U.S. RSEs that exclude the 22 blue-collar occupations. Note that by eliminating these occupations, again there is an improvement in the RSE distribution.

	Rest of U.S.		White Collar Pay	
	Dist.	Perc.	Dist.	Perc.
Less than 1%	16	23.2	72	52.6
1 and under 2%	33	47.8	51	37.2
2 and under 3%	17	24.6	9	6.6
3 and under 4%	3	4.3	5	3.6
	69		137	

Analysis

In general, a comparison between the WCP RSEs and the OCSP national and rest of the U.S. RSEs, shows that the reliability of the surveys is very similar, if not possibly better in the WCP. There are several design parameters that may explain the slight disparity.

As mentioned previously, the WCP was a national survey. For sampling, the establishments were stratified by employment and SIC; sample establishments were selected with equal probability within strata, whereas the sample selection for the OCSP national and rest of the U.S. was two-stage cluster sampling. In the first stage, the ninety primary sampling units were selected from the metropolitan areas and seventy areas were selected from the nonmetropolitan areas; the second stage was similar to the WCP single stage sampling. Therefore, the between cluster variability would increase the RSEs of the OCSP national and rest of the U.S. estimates.

A comparison of the sample sizes for private industry also is important. The sample sizes for national and the rest of the U.S. were based on white collar pay sample sizes with an additional 20% in the locality pay program to account for the cluster based design. The total number of usable establishments, that is the total sample excluding any establishments for which there was no response, was 14,353 for the national and 5,876 for the rest of the U.S. The white collar pay survey had a collected sample size of 6,625. The smaller sample size for the rest of the U.S. is a possible explanation as to why its RSEs were slightly larger.

Another possible factor in the reliability was the length of collection for the WCP versus the OCSP. The collection period for the WCP was six months; whereas the collection of all the OCSP surveys that comprised the OCSP national and rest of the U.S. used in this study spanned 21 months. Therefore, part of the variability associated with the OCSP estimates would be the normal cost of living adjustments that occur in a two year time frame.

Criticals versus AWSs

Of the 32 critical areas, RSEs were available for fourteen comparable Area Wage Surveys. These surveys had the same geographic scope, however the industrial coverage for the AWSs was a subset of the OCSP industrial scope. In general, the reliability of the estimates for the critical surveys was better than for the corresponding AWSs. Below are the one RSE distributions in the private industry for all fourteen areas combined.

	Criticals		Area Wage Surveys	
	Dist.	Perc.	Dist.	Perc.
Less than 1%	83	11.8	16	2.6
1 and under 2%	272	38.8	172	28.1
2 and under 3%	181	25.8	183	29.9
3 and under 4%	81	11.6	119	19.4
4 and under 5%	56	8.0	79	12.9
5% or more	28	4.0	43	7.0
	701		612	

In the table above, it can be seen that for the combined criticals, the majority of the RSEs fell below three percent. For the combined Area Wage Surveys more RSEs fell above 3% compared to the criticals.

The table below represents the New York critical survey compared to the New York AWS.

	New York Critical		New York Area Wage Survey	
	Dist.	Perc.	Dist.	Perc.
Less than 1%	1	1.6	1	1.6
1 and under 2%	27	44.3	19	30.2
2 and under 3%	12	19.7	17	27.0
3 and under 4%	12	19.7	7	11.1
4 and under 5%	8	13.1	12	19.0
5% or more	1	1.6	7	11.1
	61		63	

As also noted with the fourteen combined criticals, the RSEs for the New York critical were smaller than for the AWS. This can be generalized to most of the other thirteen criticals/AWSs. The explanations for the better RSEs are not as apparent as when comparing the WCP with the OCSP national and rest of the U.S. estimates, because there are more similarities than differences. For example, in the sample design of both the criticals and AWSs, all establishments with employment of 2500 or more were sampled with certainty. Also, the collection period for both survey types was approximately the same length.

Perhaps the main reason why the estimates for the criticals were more reliable than the AWSs was the differences in sample sizes. The sample sizes for all criticals were based on the corresponding sample size from the area wage survey program, with additional sample included to account for the expanded scope. Thus, for all of the criticals the sample sizes were larger than for the AWSs. Overall, the sample sizes for the criticals were almost forty percent larger than for the AWSs. For example, the sample size for the New York critical was 495 establishments compared to 372 in the AWS. Please note that these sample size comparisons are based strictly on the collected establishments in the private industry.

Another difference that might account for the better RSEs in the criticals was in the two allocation procedures. For all of these surveys there are specific occupations that are surveyed. Based on prepared job definitions, the Bureau's field staff classifies workers who are employed in the sampled establishments into the appropriate occupations. If a particular worker meets the requirements for a particular BLS-defined occupation, this is called a job match. The allocation for the AWSs was based on total employment whereas for the criticals it was based on expected number of job matches. Although the stratification variables were the same for both, there were more sampling cells in the criticals. For the criticals, these two factors (allocation based on job matches and increased number of sampling cells) enabled the sample to be concentrated in establishments where more job matches were anticipated.

CONCLUSIONS AND FURTHER RESEARCH

The new integrated program produces mean wage estimates that are as reliable as previous programs. The WCP estimates may have been slightly more reliable than the OCSP national and the rest of the U.S. estimates, but the AWS estimates were a little less reliable than the criticals. On the basis of these results, the sample sizes chosen for the locality pay surveys proved to be sufficient in producing estimates that were just as reliable as previous WCP and AWS estimates. Although the estimates were reliable, there are several changes that could be made to enhance the locality pay program.

First of all, if the collection period for the OCSP national and rest of the U.S. survey could be decreased, some of the variability caused by general cost of living differences would be eliminated. Secondly, the current allocation procedures used for the criticals could be updated with more recent expected job match data. The job match database currently being used was produced based on historical WCP data, as well as some early locality pay test data. The expected job match database could be revised to a more area-specific database compared to the national database that is currently being used. This expected job match data is presently only used for the critical surveys; it could be expanded to include all surveys in the OCSP. In addition, research could be done to determine if the current sampling cells (which are based on employment and SIC) could be defined differently so that the expected job match database would be more efficient.

The feasibility of occupational subsampling should also be investigated. If the RSEs for particular occupations are very reliable, resources for these occupations could be applied to other occupations or sampling cells where the RSEs are less reliable.

Overall, the integration has been very successful. More occupations with a broader industrial scope have been published with no major impact on the reliability of the estimates. Also state and local government data are available for the first time. With any new program, however, there is room for improvement. As the staff become better trained and more experienced, there will be more time and knowledge to dedicate to improving the program.