

Evaluation of mean wage estimates in the Industry Wage Survey program

Variances and wage distribution data provide the basis for evaluating the reliability of mean wage estimates: sampling size, worker counts, and wage dispersion were found to affect relative standard errors

PENNY L. ASBURY AND CARL BARSKY

The first annual report of the Commissioner of Labor, published in 1886, included the results of an occupational wage survey conducted by what is now the Bureau of Labor Statistics (BLS).¹ The results, taken from payroll records of 582 establishments in about 40 mostly manufacturing industries, contained daily mean wage rates by occupation, industry, and State.

Since that first report, the BLS has continued the study of occupational wages by industry. This Industry Wage Survey program now includes approximately 25 manufacturing and 15 nonmanufacturing studies, which represent a total of about 65 industries. About eight surveys per year are conducted. Most surveys are done on either a 3- or a 5-year cycle. For each survey, average (mean) wages and wage distributions for workers in selected occupations are published on a national, regional, or locality basis.

For any statistical survey program such as the Industry Wage Survey, a measure of the sampling error should be available for each mean wage estimate derived from the survey sample to provide an indication of the quality of the survey data. Sampling errors occur because the estimates are based on observations from a subset of the population rather than from the entire population. The particular sample selected for a survey is one of a large number of possible random samples of the same size that could have been selected.

The most commonly used measure of sampling errors is the variance. Accordingly, this article discusses a variance estimation procedure used in five manufacturing and two nonmanufacturing surveys from the 1985 and 1986 Industry Wage Survey program. In general, it was found that most of the variances were at the acceptable level of below 3 percent. The variances increased inversely with the sample size of the survey and with the number of workers in an occupation. However, they varied directly with the dispersion of wage rates in an occupation.

Uses of variance estimates

The purposes of calculating variances for the Industry Wage Survey program are 1) to evaluate the quality of survey data, 2) to publish information on the reliability of the survey estimates, and 3) to improve the efficiency of sample allocations. By evaluating the variances of mean wages among occupations, the BLS can improve its sampling procedures by determining the conditions under which the sample size for a given occupation or industry should be increased or decreased to provide the desired overall precision.

For the surveys discussed in this article, relative standard error, a form of variance, is used as a measure of survey reliability. A calculation of variance is converted into a relative standard error by dividing the square root of the variance by the mean wage estimate. The relative standard error is used because it measures the precision of an estimate, while eliminating the level differences caused by the different mean wage estimates among occupations. Relative standard errors permit a comparison of the reli-

Penny L. Asbury is a mathematical statistician in the Office of Compensation and Working Conditions, Bureau of Labor Statistics. Carl Barsky, formerly an economist in that office, is currently in the Bureau's Division of Systems Design.

ability of mean wage estimates between different occupations or geographic areas.

For example in the Industry Wage Survey of hospitals, the mean wage for the occupation head nurse can be compared across all metropolitan areas studied. In Oakland, CA, the mean wage was \$17.53 an hour and in Buffalo-Niagara Falls, NY, it was \$11.89 an hour. The relative standard errors were 0.94 for Oakland, and 0.92 for Buffalo-Niagara Falls. The relative standard errors show that for both areas the mean wage estimates, although different, are equally reliable. When comparing two estimates, a smaller relative standard error indicates greater precision.

The estimated relative standard errors can also be used to calculate a 95-percent confidence interval around the mean wage estimate. A 95-percent confidence interval means that if similar samples were repeatedly drawn from the same population, and estimates of the mean wage and its relative standard error were computed for each sample, then the true population mean would be included in the confidence interval for approximately 95 percent of these samples.

A 95-percent confidence interval has a lower limit equal to the estimated mean wage minus 2 times the relative standard error times the estimated mean wage, and an upper limit equal to the estimated mean wage plus 2 times the relative standard error times the estimated mean wage. For example, the nationwide estimated mean wage for production workers in the survey of the petroleum refining industry was \$14.20 in 1986, with a relative standard error of 0.23 percent. Therefore, a 95-percent confidence interval for the estimate is from \$14.13 to \$14.27. (The lower confidence limit is \$14.20 minus 2 times 0.0023 times \$14.20, or \$14.20 minus \$0.07. The upper limit is \$14.20 plus \$0.07.)

Characteristics of evaluated surveys

The surveys covered by the variance estimation procedure discussed in this article were mostly in manufacturing: cotton and manmade textiles, synthetic fibers, petroleum refining, industrial chemicals, and glassware. There were more than 100 establishments in the sample for all manufacturing surveys except that for synthetic fibers which, because of the industry's size, included only 37 establishments. The surveys provided mean wage estimates on a national or regional basis with industrial chemicals and cotton and manmade textiles also providing some locality estimates.

The two nonmanufacturing surveys, hospitals and nursing homes, had sample sizes of around 500 establishments each, and provided estimates for approximately two dozen metropolitan areas.²

These seven surveys were chosen to evaluate the general Industry Wage Survey program because of their varying

degrees of statistical complexity. The hospitals and nursing homes surveys involved simple sample designs which provided mean wage estimates only by locality. More complex sample designs, such as those used in the surveys of the cotton and manmade textiles and industrial chemicals industries, provided estimates not only at the locality level, but also at regional and national levels. The industrial chemicals survey also produced separate estimates for the inorganic and organic chemicals subclassifications.

Because sample designs vary by survey, the variance estimation procedure must be modified for each survey in the Industry Wage Survey program. For locality surveys, the procedure is straightforward. However, for surveys involving national, regional, and locality estimates, the procedure must be adapted for each level of estimation.

Sampling design

The variance estimation procedure used to compute relative standard errors for any survey depends on the sampling design of the survey and the estimator. For sampling, the establishments in the Industry Wage Survey are separated by the characteristics associated with wage differences, such as geography and number of employees. Then, a simple random sample is chosen from each group (or cell) of establishments with similar characteristics. The assumption is that occupational wages and benefits tend to be similar among establishments with similar characteristics.

The number of sample establishments in each cell chosen for a survey is based on the proportion of employment in that cell to the employment of establishments within the scope of the industry. In practice, because the sampling design assumes that variance is proportional to the number of workers in an establishment, the usual consequence of this is that a cell which contains 10 percent of the total industry employment is allocated approximately 10 percent of the total sample establishments. There are two additional constraints that are imposed on the sample allocation procedure to reduce variance and to ensure minimum bias in sampling and nonresponse adjustment procedures:³

- 1) All establishments with 2,500 employees or more are included in a survey sample; and
- 2) Minimum sample sizes are required for each cell based on the total number of establishments in the cell.⁴

Industry Wage Survey samples would ideally be designed so that estimates of average wages have relative standard errors no greater than 7.5 percent. However, the Unemployment Insurance file, which serves as the source for the survey universe of establishments in an industry, does not include any information on wages. (A universe is a list of all eligible establishments from which a sample is chosen.) Employment size is the only measure

of establishment characteristics available from the Unemployment Insurance file. Therefore, sample size and sample allocation for the surveys have been determined under the requirement that estimates of total employment have relative standard errors no greater than 7.5 percent. The validity of this approach to Industry Wage Survey sample design rests on the assumptions that wages are less variable than establishment size in terms of number of employees and that the number of workers in the occupations studied is directly proportional to establishment size.

As the relative standard errors are calculated for the different Industry Wage Surveys, they will be compared from occupation to occupation to determine whether the sampling design requirements are fulfilled. After evaluation, it may be determined that some occupations will need more observations in future surveys to obtain the required precision, while the number of observations may be decreased for others.

Variance estimation procedure

For the surveys evaluated in this article two variance estimation procedures were considered. The first was a replication technique.⁵ In this procedure, the survey is divided into subsamples (replicates) in accordance with the sampling design, and estimates of mean wages are computed for each. Then, the sample variance among the several mean wage estimates is computed. This is a relatively simple procedure, and with large sample sizes produces an accurate estimate of variance.

The estimation procedure which was actually used in calculating the variances is an approximation of the formulas used to produce the survey mean wage estimates.⁶ Although it is more involved than the replication technique, it provides more reliable estimates of variances for the wage surveys which have relatively small sample sizes.

Implementing the variance estimation procedure is difficult because it must be modified for each survey. Any sample cell with only one establishment must be combined with another cell with similar characteristics, because the procedure does not allow for the computation of a relative standard error for a cell with one establishment.

Each survey also must be evaluated for sampling areas that overlap. For example, in industrial chemicals, the data used to produce locality estimates for Philadelphia, Newark, and Buffalo must be combined with the data for the rest of the Middle Atlantic region to compute regional estimates.

Relative standard errors are calculated on mean wage estimates for each occupation in each published tabulation. In the industrial chemicals survey, for example, wage estimates are published not only for the overall industrial chemicals classification, but also for the organic and inorganic chemicals industries. These figures include

estimates for the Nation, and for nine economic regions. Estimates also are published for the overall industrial chemicals classification for eight localities of industry concentration. The 35 occupations for each industry sector and geographic tabulation in the survey result in 1,330 possible mean wage estimates for which relative standard errors can be computed.

In the less complicated nursing homes industry survey, estimates are published for three categories (all workers, full-time, and part-time) in 15 professional and technical occupations in 22 localities for a possible total of 990 mean wage estimates. Because there are no overlapping areas, the relative standard errors are easier to compute.

Analysis of relative standard errors

For the surveys studied, 85 out of the possible 120 locality, regional, and national wage tabulations were analyzed.⁷ As the following tabulation shows, of those relative standard errors that were calculated, most are under 3 percent.⁸

<i>Relative standard error</i>	<i>Percentage of occupations</i>
Less than 1 percent	27
1 and under 2 percent	36
2 and under 3 percent	17
3 percent or more	<u>19</u>
Total	100*
Number of occupations evaluated	2,934

*Due to rounding, sums of individual items do not total 100.

In general, the relative standard errors for national estimates are lower than those for regional estimates which, in turn, are lower than those for locality estimates. Note from the tabulation below how the quality of the estimates improves as geographic areas become larger:

<i>Relative standard error</i>	<i>National</i>	<i>Regional</i>	<i>Locality</i>
Less than 1 percent	45	35	21
1 and under 2 percent	30	28	41
2 and under 3 percent	11	16	19
3 percent or more	<u>14</u>	<u>20</u>	<u>19</u>
Total	100	100*	100
Number of occupations	248	916	1,770

*Due to rounding, sums of individual items do not total 100.

This pattern occurs because the relative standard error of an estimate generally varies inversely with the sample size of the survey. The national estimates have a larger number of establishments in their samples and smaller relative standard errors than the regional or locality estimates from the same survey. Because the hospitals and nursing home surveys are designed to obtain only locality estimates, their estimates are not as reliable as the other surveys, which provide mostly regional and national estimates.

The relative standard error can also vary inversely with the number of workers sampled in an occupation. This explains why the national mean wage estimates for occupations with large worker counts have smaller relative standard errors than the regional or locality wage estimates with their smaller worker counts. However, because of the sampling design it should be noted that relative standard errors are calculated on establishment wage means and not on wages for individual workers.

An inverse relationship was also found between relative standard error and occupation with different employment level, as the tabulation below shows:

Relative standard error	Percent of occupations with employment of—				
	Less than 100	100 and under 250	250 and under 500	500 and under 1,000	1,000 and over
Less than 1 percent	22	24	27	32	44
1 and under 2 percent	25	34	46	53	46
2 and under 3 percent	21	23	17	8	7
3 percent or more	33	18	10	7	2
Total	100*	100	100	100	100*
Number of occupations ...	1,158	629	412	282	453

*Due to rounding, the sums of individual items do not total 100.

Nine-tenths of the occupations with 1,000 workers or more had relative standard errors of less than 2 percent, whereas slightly more than half of the occupations with fewer than 100 workers had relative standard errors exceeding 2 percent. For example, in the container segment of the glassware survey for the United States, the occupation batch mixer has 153 workers and a relative standard error of 1.50 percent, while mold metal maker, with 1,280 workers, has a relative standard error of 0.25 percent.

Thus, when an occupation has a large number of workers, the relative standard error of the estimate is lower. The "all production worker" estimate in manufacturing surveys is another good example. Because this broad employment category includes all production workers from each region, State, or locality, it has the largest number of workers contributing to a mean wage estimate, and should have a small relative standard error.

Of the 51 relative standard error estimates for the all production worker level in the five manufacturing surveys, half are less than 1 percent. Nine-tenths of these relative standard errors are under 2 percent. Similarly, the smallest relative standard errors in the hospitals and nursing homes surveys are in the occupations, such as licensed practical nurse and general duty nurse, which have the largest worker counts.

Relative standard errors are also directly related to the dispersion of wage rates in an occupation. A mean wage estimate for an occupation with a large dispersion of wage rates is more likely to have a large relative standard error

than an estimate for an occupation with less wage dispersion, unless the sample is extremely large.

To illustrate, in the industrial chemicals survey, relative standard errors are larger for the occupations in organic chemicals than for those in inorganic chemicals. A comparison is presented in the following tabulation:

Relative standard error	Percent of occupations	
	Inorganic chemicals	Organic chemicals
Less than 1 percent	38	31
1 and under 2 percent	31	22
2 and under 3 percent	15	17
3 percent or more	15	30
Total	100*	100
Number of occupations	124	132

*Due to rounding, sums of individual items do not total 100.

Organic chemicals has a wider variety of processes which creates a wider dispersion in occupational wage rates. Conversely, inorganic chemicals wages are less dispersed not only because the industry has few processes but also because it is highly unionized.

Another highly unionized industry, petroleum refining, has a narrow dispersion of wages and consequently the smallest relative standard errors of all industry surveys studied. Almost three-fourths of the relative standard errors for occupations in the petroleum refining survey are under 1 percent.

Occupations or industries with wide wage dispersions require larger sample sizes to generate acceptable relative standard errors. Conversely, selected occupational sampling (collecting wage data for particular selected occupations from only a subset of the sample) should be possible for those occupations with large worker counts and narrowly dispersed wage rates. A variance estimation procedure is necessary to identify the situations in which this is possible.

To illustrate this point, the occupation general duty nurse in the hospitals survey has comparatively small relative standard errors for mean wages in all areas surveyed, ranging from 0.54 percent to 1.01 percent. By contrast, if half of the sampled establishments were used for this occupation, then these relative standard errors would increase to only 0.57 percent and 1.85 percent. Thus, general duty nurses in the hospitals survey would be a valid candidate for selected occupational sampling.

Wage distribution as assessment tool

The published releases and bulletins of the Industry Wage Survey contain data on the distribution of workers by straight-time hourly wages in selected occupations. These distributions can also be used to assess the reliability of survey data. Distributions around the mean wage rates show the dispersion of the data that relative stan-

dard errors measure. A small relative standard error reflects a small spread in the distribution of wages, or a large number of workers in the occupation, or both.

Relative standard errors provide convenient, reliable measures of variability. However, the published wage distribution tables can be used to explain the relative standard errors and to present more information as well. The wage distribution tables include not only the lowest and highest wage rates surveyed, but also the concentration of observations in between the extremes. The tables also provide estimates of the number of establishments and employment within the survey coverage along with the actual number of establishments in the survey sample.

Survey sample sizes give an additional indication of the quality of a mean wage estimate. Reliability of survey data is related to the sampling ratio. Thus, an estimate derived from 50 workers in a sample of 7 out of 8 establishments will probably be more accurate than an estimate calculated from 250 workers in a sample of 40 out of 80 establishments.

The effect of the distribution of wage rates on the variance calculation is evident for janitors in the petroleum refining survey. Two regions, Midwest I and Midwest II,⁹ had similar sample sizes and sampling ratios. The wage spread in the Midwest II region, however, was larger than that in Midwest I. The larger relative standard error of 2.03 percent in the Midwest II region, compared to a relative standard error of 0.42 percent in Midwest I, is due to the larger wage spread. (See table 1.)

Occupations that have workers clustered at two or more points in the distribution usually have large relative standard errors. The mean wage falls between and poorly represents the wage clusters. In this situation, the mean, by itself, does not provide a clear indication of where wage rates are concentrated.

An example of this occurs for the occupation chemical operator II in the industrial chemicals survey in Newark, NJ. (See table 2.) In this locality, the wage spread for the occupation of electrician was more concentrated, with a

Table 2. Wage distributions for selected occupations in industrial chemicals in Newark, NJ, June 1986

Chemical wage distribution	Chemical operator II	Electrician
Number of workers.....	207	19
Mean hourly wages.....	\$13.46	\$12.65
Relative standard error (percent).....	.82	.25
Percent of workers receiving straight-time hourly wages of—		
Less than \$11.75.....	11	—
\$11.75 to \$12.00.....	—	37
12.00 to 12.25.....	37	11
12.25 to 12.50.....	—	32
12.50 to 12.75.....	—	—
12.75 to 13.00.....	—	—
13.00 to 13.25.....	—	—
13.25 to 13.50.....	—	—
13.50 to 13.75.....	12	—
13.75 to 14.00.....	—	—
14.00 to 14.25.....	—	—
14.25 to 14.50.....	2	—
14.50 to 15.00.....	5	21
15.00 to 15.50.....	3	—
15.50 to 16.00.....	30	—

large proportion of workers falling in a single wage interval, from \$11.75 to \$12.50. As expected, chemical operators II, with a concentration of wages at two levels, \$12 to \$12.25 and \$15.50 to \$16, had a larger relative standard error (0.82 percent) than electricians (0.25 percent).

In the cotton and manmade textile Industry Wage Survey, 7 out of 11 establishments were surveyed in Burlington, NC. The mean wage for the 202 loom fixers employed by these firms was \$8.65 an hour with a relative standard error of 0.73 percent. In Georgia, 40 out of 110 establishments were surveyed. The wages of 895 workers employed as loom fixers were \$8.29 an hour with a relative standard error of 1.32 percent. The relative standard error for Burlington is smaller for two reasons: the high sampling ratio and the greater concentration of the wage data. (See table 3.)

As discussed previously, worker counts also are related to the quality of the survey estimates. In the hospitals survey, the occupation of general duty nurse in Boston and Milwaukee has similar sample sizes and similar wage dispersions, but the relative standard error was 0.89 percent in Boston and 1.01 percent in Milwaukee. The slightly smaller relative standard error in Boston is due partly to the larger number of workers surveyed—8,260, compared to 2,680 in Milwaukee.

One cautionary note is necessary on the use of wage distribution data. As indicated earlier, relative standard errors are calculated on establishment wage means, not on wages for individual workers depicted in the wage distributions. Thus, a wide range of worker wages does not always yield a large relative standard error, even if the distribution is wide within each establishment. However, if the distribution of wages within each establishment is closely concentrated, but the establishment mean wage

Table 1. Wage distributions for janitors in petroleum refining, selected regions, June 1985

Wage distribution	Midwest I	Midwest II
Number of workers.....	22	20
Mean hourly wages.....	\$13.20	\$12.00
Relative standard error (percent).....	.42	2.3
Percent of workers receiving straight-time hourly wages of—		
\$8.50 to \$9.00.....	—	10
9.00 to 9.50.....	—	—
9.50 to 10.00.....	—	—
10.50 to 11.00.....	—	—
11.00 to 11.50.....	—	—
11.50 to 12.00.....	—	40
12.00 to 12.50.....	—	5
12.50 to 13.00.....	18	40
13.00 to 13.50.....	55	5
13.50 to 14.00.....	27	—

varies substantially among establishments, a large relative standard error will result.

In the industrial chemicals survey, for example, the wages of the occupation instrument repairer range from \$11 to over \$20 with a mean of \$15.64. However, the relative standard error is only 1.07 percent. This comparatively small relative standard error results from establishment means which are closely concentrated, not from the actual wages of the repairers.

Future possibilities

Although the variance estimation procedure has been successfully applied in a variety of Industry Wage Surveys, there are further projects that need to be undertaken. The relative standard errors and variance calculations could be programmed into the occupational wage survey computer system so that they can be published concurrently with the survey results.

Because of the number of different estimates produced in each Industry Wage Survey (and the sample design differences between surveys), computing and publishing the relative standard errors on a regular basis will require resource and publication trade-offs. The publication alternatives are to 1) provide the relative standard errors for all survey mean estimates; 2) provide a graph of computed generalized variances (a technique useful for surveys which publish a large amount of data); 3) provide frequency table distributions of the relative standard errors associated with the occupation means; or 4) publish only the mean wage estimates of those occupations which meet a specified precision.

The relative standard errors can also be used to evaluate and improve the efficiency of the Industry Wage

Table 3. Wage distribution in cotton and manmade textiles for loom fixers, selected areas, June 1985

Textile wage distribution	Burlington, nc	Georgia
Number of workers.....	202	895
Mean hourly wages.....	\$8.65	\$8.29
Relative standard error (percent).....	.73	1.32
Percent of workers receiving straight-time hourly wages of—		
Less than \$6.75.....	—	3
\$6.75 to \$7.00.....	—	5
7.00 to 7.25.....	—	5
7.25 to 7.50.....	—	1
7.50 to 7.75.....	—	—
7.75 to 8.00.....	—	11
8.00 to 8.25.....	—	14
8.25 to 8.50.....	12	11
8.50 to 8.75.....	72	40
8.75 to 9.00.....	12	1
9.00 and over.....	4	8

Survey sample allocation procedure. By comparing the relative standard errors among the mean wage estimates for the different occupations in a survey, the Bureau of Labor Statistics will be able to evaluate the sample sizes for each survey and adjust them accordingly. It might be possible to sample selected occupations to reduce respondent burden when relative standard errors indicate that this is possible or revise the occupation list if the relative standard errors indicate a problem.

Finally, if possible, the relative standard errors will be computed using a replication technique. Computer simulation of this approach might be compared to the results obtained by the current procedure to determine if the results are similar. If the replication method gives comparable results, it might be chosen as a more efficient production method to obtain the relative standard error data. □

FOOTNOTES

¹For more detailed information on the early years of the Industry Wage Survey program, see H.M. Douty, "A century of wage statistics: the BLS contribution," *Monthly Labor Review*, November 1984, pp. 16-28.

²For detailed accounts of these surveys, see *Industry Wage Survey: Petroleum Refining, June 1985*, Bulletin 2255; *Industry Wage Survey: Textile Mills, June 1985*, Bulletin 2265; *Industry Wage Survey: Hospitals, August 1985*, Bulletin 2273; *Industry Wage Survey: Synthetic Fibers, September 1985*, Bulletin 2268; *Industry Wage Survey: Nursing and Personal Care, September 1985*, Bulletin 2275; *Industry Wage Survey: Pressed or Blown Glass and Glassware, June 1986*, Bulletin 2286; *Industry Wage Survey: Industrial Chemicals, June 1986*, Bulletin 2287 (Bureau of Labor Statistics).

³Nonresponse adjustment involves reweighting for those establishments for which no data were obtained.

⁴For a more complete description of the sampling design of the IWS program, see *BLS Handbook of Methods*, Bulletin 2285 (Bureau of Labor Statistics, 1988), ch. 6, pp. 41-48.

⁵Kirk M. Wolter, *Introduction to Variance Estimation* (New York, Springer-Verlag, 1985).

⁶William G. Cochran, *Sampling Techniques* (New York, John Wiley & Sons, 1977).

⁷Relative standard errors were not computed for all published tables, because the approximation used to develop the relative standard errors is only valid for sample sizes of 30 or more. For universes with smaller sample sizes and large sampling fractions (ratio of sample establishments to universe establishments), the relative standard errors can be used to evaluate the relative quality between two estimates. Confidence intervals, however, have little meaning because the normal theory assumptions on which they are based usually are not valid. Also, when the sample consists of almost all of the units in the universe, the relative standard errors are essentially meaningless because there is minimal variability associated with sampling.

⁸These relative standard error limits were chosen to correspond with the frequency table of the Bureau's National Survey of Professional, Administrative, Technical, and Clerical Pay (PATC). Occupational levels in the PATC are published only if their relative standard errors are smaller than 5 percent.

⁹The region Midwest I includes Illinois, Indiana, Kentucky, Michigan, Ohio, and Tennessee. Midwest II includes Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, Oklahoma, South Dakota, and Wisconsin.