Pay Ratio: “Other Reasonable Methods”

A Simulation Approach to Estimate Median Employee Pay

NASPP Connecticut / Boston Chapter Conference
July 2014

Andy Restaino
andy.restaino@techcompadvisors.com
(516) 804-6135

Technical Compensation Advisors, Inc.

Overview of Proposed Pay Ratio Rules

- The SEC proposed adding Item 402(u) to the executive compensation disclosure rules to implement Section 953(b) of Dodd-Frank
  - Note that this is a different section than the CD&A (i.e., ratio might be disclosed elsewhere)
- "Total Compensation" for the CEO and the median employee calculated using Summary Compensation Table rules for NEOs
- All employees need to be considered in calculating the median including full-time, part-time, seasonal or temporary workers as well as non-US employees
- The SEC is proposing that companies comply with the rule for the first fiscal year after the effective date of the rule
  - For calendar-year companies, likely to be 2016 proxy disclosure for 2015
- To identify the “median employee,” proposed Instruction 2 to Item 402(u) would permit the use of:
  - Reasonable estimates to identify median or to calculate elements of total compensation for median
  - Statistical sampling or “other reasonable methods”
  - Other consistently applied compensation measures to identify the “median employee”
    - Once “median employee” is identified, total compensation must be calculated in accordance with Summary Compensation Table rules

Statistical Sampling or "Other Reasonable Methods"

- The preamble of the proposed rules places considerable emphasis on statistical sampling (several pages), but provides virtually no discussion of "other reasonable methods"
  - Brief mention of "statistical inference or numerically solved median estimate," but only in the context of conducting a stratified sampling approach
- Statistical sampling may be a viable approach for a number of companies but for others, it might not be as cheap and easy as the SEC implies
  - Extracting sample data from multiple payroll systems (e.g., manually, computer program, etc.)
  - Determination of sample size and confidence intervals
  - Effort spent sampling might exceed effort spent arraying data and finding median directly
    - Sampling requires ability to access all of the data anyway
  - Data for certain units or segments may not be available (e.g., non-US unit run independently) which makes sampling impossible
    - But, could make inferences about variance if descriptive statistics provided for unit
- Other reasonable methods might include:
  - Statistical inference
    - SEC cites paper that discusses the use of weighted-average medians, but only in the context of stratified cluster sampling
  - Simulation
    - Not mentioned proposed rules

Simulation Methodology

- Proposed simulation assumes compensation data follow a “lognormal distribution”
  - Common assumption with academic and BLS research (including SEC citations in proposed rules)
  - Data can’t be below $0 and are assumed to be skewed to the right with mean higher than median
  - Adjust compensation data by taking natural log (i.e., log base e or LN(x) in Excel)
    - Ranking of data is exactly the same before and after data are adjusted
    - Adjusted data assumed to follow a normal distribution (bell curve)
    - Normal distribution has desirable mathematical properties (e.g., symmetrical with mean = median)
    - Properties of normal distribution are well-known and easily applied to create simulation approach
      - Other distributions would be more complicated
- Once a normal distribution can be assumed, simulation only requires:
  - Mean = μ
  - Standard deviation = σ
  - Stochastic factor from random number generator (e.g., NORM.S.INV(RAND()) in Excel) = z

Simulation Methodology (continued)

- If quartile data are provided for a business unit, then can take the following steps:
  1. Take the natural log of each quartile
     - These data are now assumed to be normally distributed with mean = median
  2. Log-adjusted median data is also mean (μ) of normal distribution
  3. Algebraically solve for the standard deviation
     - Since log-adjusted data are assumed to be normally distributed (i.e., symmetrical bell-curve), the statistical properties of a normal distribution can be applied
       - These properties are well-known and can be found in virtually all statistics text books
     - 25^{th} percentile is -0.675 standard deviations below the mean (median) and 75^{th} percentile is +0.675 standard deviation above the mean (median)
       - Solve for standard deviation from these quartiles
       - Take average for standard deviation input (σ)
       - Can use other percentiles if available (e.g., the 16^{th} and 84^{th} percentiles are ± 1 standard deviation from mean)
  4. Come up with a random number between 0 and 1 (e.g., RAND() function in Excel) and convert to a standard normal random variable (z), (e.g., use table from textbook or NORM.S.INV in Excel)
  5. Simulate log-adjusted compensation = $\mu + z\,\sigma$
  6. Repeat for the number of remaining employees in the business unit
- Follow the above process for each business unit where data are lacking
  - Approach implicitly assumes no correlation between different units
- Combine all actual and simulated data into one series and take median
- Take the anti-log (i.e., e^x) of the result to find median pay (and other percentiles)

Simulation Example

- Quartile data were provided for Global population as well as US and non-US populations
- Test approach by trying to replicate median of global data by using only US and non-US quartile data
- Data provided below (actual and log-adjusted):

Salary ($ US)			
	Global	*Non-US*	*US*
Percentile	*n = 50,000*	*n = 31,000*	*n = 19,000*
25th	$ 11,250	$ 5,750	$ 35,750
50th	$ 34,250	$ 17,250	$ 51,500
75th	$ 56,500	$ 41,000	$ 77,750

ln(Salary)			
	Global	*Non-US*	*US*
Percentile	*n = 50,000*	*n = 31,000*	*n = 19,000*
25th	9.33	8.66	10.48
50th	10.44	9.76	10.85
75th	10.94	10.62	11.26

- Mean for Non-US is 9.76 and mean for US is 10.85
- Solve for standard deviations:
  - Non-US: [(8.66 - 9.76) ÷ -0.675 + (10.62 – 9.76) ÷ 0.675] ÷ 2 = [1.63 + 1.27] ÷ 2 = 1.45
  - US: [(10.48 - 10.85) ÷ -0.675 + (11.26 – 10.85) ÷ 0.675] ÷ 2 = [0.55 + 0.61] ÷ 2 = 0.58*
- Use Excel RAND() and NORM.S.INV functions and simulate data point
  - Example: NORM.S.INV(RAND()) = 0.4232 → for US data point, 10.85 + 0.4232 x 0.58 = 11.10; $e^{11.10}$ = $66,171
- Simulate 31,000 data points for non-US and 19,000 for US, combine into one series and determine median; repeat many times and average all results

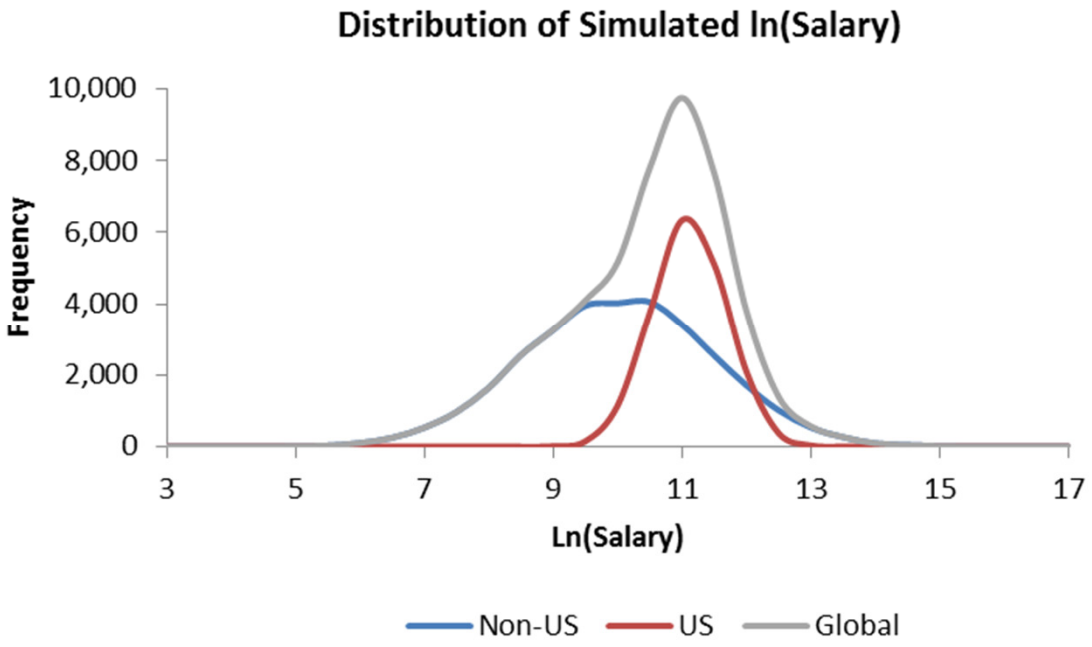
* Note: Might be more appropriate to take average of variance rather than average of standard deviations; difference is negligible in this example

Simulation Example (continued)

- Results from one simulation is provided as follows:

ln(Salary)			
	Global	*Non-US*	*US*
Percentile	*n = 50,000*	*n = 31,000*	*n = 19,000*
25th	9.39	8.77	10.47
50th	10.42	9.77	10.86
75th	11.07	10.76	11.23

Salary ($ US)			
	Global	*Non-US*	*US*
Percentile	*n = 50,000*	*n = 31,000*	*n = 19,000*
25th	$ 12,009	$ 6,428	$ 35,286
50th	$ 33,460	$ 17,544	$ 51,807
75th	$ 64,283	$ 46,906	$ 75,592



Simulation Example (continued)

- The following table compares the results from this one simulation to actual (the difference between the two are displayed):

Salary ($ US)			
	Global	*Non-US*	*US*
Percentile	*n = 50,000*	*n = 31,000*	*n = 19,000*
25th	$ 759	$ 678	$ (464)
50th	$ (790)	$ 294	$ 307
75th	$ 7,783	$ 5,906	$ (2,158)

- Note that this was just one simulation
  - The process can be repeated many (e.g., 10,000) times
  - Take an average of the results to get a more robust answer

Issues to Consider

- This approach could be useful when trying to integrate multiple data series
  - Dispersion for a business unit could influence position (ranking) of data in the global population
  - Actual data are less important than the ranking when finding a median
    - Unlike a mean, magnitude of data doesn't influence the middle-point
- Approach heavily relies on assumption that pay data follow a lognormal distribution
  - Actual distribution might differ (e.g., note that quartiles in example showed some skewness)
  - SEC proposal cites BLS and academic studies that rely on this assumption
  - Proposed rules do not appear to prohibit reasonable assumptions like this, but it does criticize variances that are "generally assumed"
    - Using descriptive statistics that are specific to a company or unit addresses this concern
    - Specific guidance form the SEC would help to quell any potential criticism
  - Simulation based on other distributions might be developed (if these can be identified)
    - Less practical (more complicated) – need to incorporate skewness, kurtosis into model
- The example shows evidence that simulation would be viable, but not proof
  - The answer (i.e., the overall median) is already known
- The approach relies on obtaining some measure(s) of dispersion around the median
  - If only median data are provided, then deriving input for the standard deviation would be impossible
    - Simulation approach would need to rely on additional assumptions
    - Other approaches may be more palatable
- Simulation demonstrated does not consider correlation between different groups of data
  - Implicit assumption is that there is no correlation between data sets
  - Can incorporate correlation into simulation

Recommendations to the SEC

- If the SEC goes forward with issuing final rules, hopefully it will consider the following:
  - Alternatives for determining pay ratio should focus on practicality and cost reduction, not on spurious precision
    - Inherent tradeoff between incremental cost and incremental benefit of more "precision" (e.g., would an investor's actions change if the ratio were higher/lower by 50?)
  - Explicitly permit (1) use of lognormal assumption and (2) use of formulaic or numerical approaches to estimate median pay (e.g., simulation)
    - On the one hand, proposed rules rely heavily on studies that make lognormal assumption (footnote 204), but criticizes the assumption elsewhere (footnote 83 – but only in the context of determining sample size)
      - Using descriptive statistics to approximate company- or segment-specific compensation variance should address this criticism
  - Allow companies to provide a reasonable range of ratios based on reasonable assumptions
    - For example, if using salary to identify "median employee:"
      - Assuming $0 for all other elements would provide lowest value (and highest, most conservative ratio)
      - Assuming "median employee's" age and service qualifies him for early retirement provides highest pension change
    - If a range of outcomes is deemed acceptable, then any effort spent being more "accurate" would be a waste of shareholders money
- Full comment letter sent to SEC: http://www.sec.gov/comments/s7-07-13/s70713-570.pdf