

Attachment K
OSH Program Survey Sample Size Requirements

Aim-1 - Research Question 2 – Factor Analysis. Norman and Streiner (1994) suggest as a rule of thumb for factor analysis that there be 20 observations per item being considered in the factor analysis. Here, researchers propose examining 143 possible items for factor analysis, including OSH organizational metrics (131 possible items) and one measure (13 items) of social desirability. Following Norman and Streiner’s suggestion, information is needed from 2860 firms to conduct the factor analysis.

Aim-2 - Research Question 5 – Predictive Regression. The primary interest is observing whether organizational and management metrics predict injury rates. Sample size requirements are examined in two ways: 1) what sample size is required to detect a reasonable difference in injury rates between two groups of firms, 2) what sample size is required to fit multiple regression models with injury rate as the dependent variable. Based on BLS data the overall incidence rate is approximately 6.49 per 100 FTE with firm level claim rates ranging from 0 to 35 per 100 FTE for small firms and from 0 to 12 per 100 FTE for large firms. Based on this information, sample size requirements are examined under two sets of assumptions about firm-level variation in claim rate; standard deviation of 6 (applicable to large firms) and standard deviation of 12 (applicable to small firms). Using this information, the number of firms is calculated required in each of two groups to detect a difference of from 0.75 to 2.0 in the claim rate, with standard deviation of claim rate as either 6 or 12. Type 1 error rate was set at 0.05 and type 2 error rate was set at 0.20.

Standard Deviation of Claim Rate	Difference in Claim Rate to Detect	Sample Size (# Firms) Required in Each of Two Arms
6	0.5	2261
6	1.0	566
12	1.0	2261

If approximately 3700 firms (**Appendix 7**) all together are sampled, there will be sufficient numbers to detect reasonable differences – e.g. a difference of 0.5 to 1 claim per FTE - between two groups of firms. For regression modeling, rules of thumb (Harrell 2001) suggest there should be 20 observations (here, firms) for each term to be considered in model building. Here, there are 30 subscales to be considered representing organizational policies and practices and occupational health and safety management systems. In addition, researchers will control for sector (2 categories, 2 df), firm size (four categories, 3 df), and downsizing (yes or no, 1 df). All together that totals to 37 beta coefficients to estimate, and so the rule of thumb suggests a requirement for at least 740 firms.

Aim-3 - Research Question 6 – Who is the Best Informant. Based on approach described by Kraemer and Korner (1976) and by Kraemer (1976), if it is wished to establish an inter-rater reliability, estimated by an intraclass correlation coefficient ρ_0 , is larger than 0.8 with type 1 error $\alpha=0.05$ and type 2 error $\beta=0.20$ when the actual reliability was equal to $\rho_1=0.9$, then the sample size required should be at least n where:

$$n \geq 2 + \frac{4(z_\alpha - z_\beta)^2}{\tau^2} \quad \text{and} \quad \tau = \ln\left(\frac{(1 - \rho_1)}{(1 - \rho_0)}\right) - \ln\left(\frac{(1 + \rho_1)}{(1 + \rho_0)}\right)$$

Substituting values of ρ_0 , ρ_1 , α and β into the expression above gives $n=58.2$, so we will strive for an inter-rater-reliability sample of 60 firms.