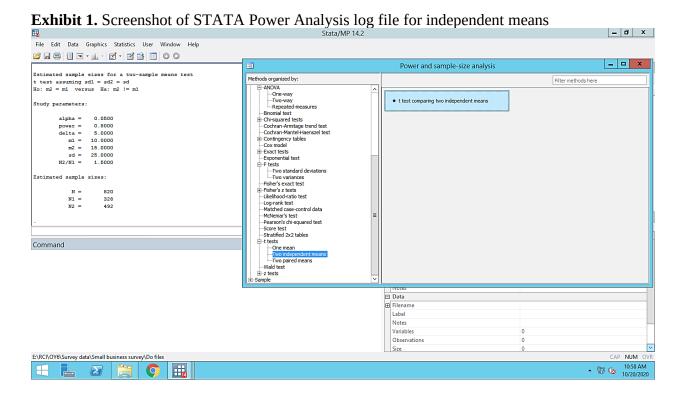**Appendix D-1: Power Analysis**

This appendix describes the rational and assumptions used to conduct the power analysis to determine the sample size needed to detect significant differences between two groups of interest (e.g., borrowers with two different delivery modes of participation in technical assistance or borrowers with and without a certain business characteristic) in the outcome measure of the change in the number of employees over time.

To address the research questions about the effect of the business characteristics and technical assistance on business growth over time, the evaluation will conduct OLS regressions for the significance of the changes in revenue and employment from the loan initiation year to 2019 as a function of the key predictors, while controlling for the relevant covariates. The key predictors are expected to be dichotomous or dummy-coded variables for the business characteristics and various dimensions of training (e.g. the mode of training). The significance of each predictor will be examined over and above the effect of all other predictors and covariates in the model. However, it is unknown whether a partial correlation of a key predictor in the context of the regression model will be higher or lower than its bi-variate correlation to the outcome. Therefore, we conducted the power analysis for a two-sample means test, which is a more conservative approach than the power analysis for the overall correlation coefficient (R^2) of the regression model. The power analysis for the regression is used to determine the sample size required for the F test for the overall R^2 in the regression model. The F-test of the overall significance indicates whether the regression model provides a better fit to the data than a model that contains no independent variables. However, the goal of the analyses to be performed is to detect the significance of the key predictors in the model and not to test the significance of the proportion of the variance in the outcome accounted for by all predictors in the model. Therefore, the power analysis for a two-sample means test is applicable.

For this power analysis, a set of assumptions was used based on the results of a previous small business evaluation conducted by Optimal for the SBA. This study provided descriptive results for the change in the number of employees over the three-year period by such business characteristics as women- and minority-owned status, business certifications and designations, and being in business for at least three years. These variables are similar to those that will be used as predictors in the regression models to be conducted by this evaluation to detect the significant difference between two groups of interest (e.g., borrowers with two different delivery modes of participation in technical assistance or borrowers with and without a certain business characteristic) in the outcome measure of the change in the number of employees over time. We also assumed that the two groups of interest (e.g. borrowers with two different delivery modes of technical assistance) will not be of equal size and the ratio of 1.5 represents an expected average ratio of the two groups for the dichotomous predictors to be used in the regression models.

The result of this power analysis for a two-sample means test indicated that the expected sample size for the borrower survey is 820, based on the 80 percent power, an alpha level of .05, the mean difference of 5, the standard deviation of 25, the ratio of the two groups of 1.5, and the effect size of .20 (see Exhibit 1 below).

Exhibit 1. Screenshot of STATA Power Analysis log file for independent means

STATA Power Analysis log file for independent means

```
Power two-means 10 15, sd (25) nratio (1.5)

Performing iteration ...

Estimated sample sizes for a two-sample means test
t test assuming sd1 = sd2 = sd
Ho: m2 = m1 versus Ha: m2! = m1

Study parameters:

        alpha =     0.0500
        power =     0.8000
        delta =     5.0000
           m1 =    10.0000
           m2 =    15.0000
           sd =    25.0000
        N2/N1 =     1.5000


Estimated sample sizes:

            N =        820
           N1 =        328
           N2 =        492
```