

SUPPORTING STATEMENT FOR REQUEST FOR EMERGENCY ACTION–PART B
 UNITED STATES INTERNATIONAL TRADE COMMISSION QUESTIONNAIRE
SMALL AND MEDIUM-SIZED ENTERPRISES: CHARACTERISTICS AND PERFORMANCE

B. Collections of Information Employing Statistical Methods

1. Respondent universe and sampling strata

The Commission will utilize the Orbis database, a proprietary database, to develop a sample of business firms. The Commission will sample firms in the manufacturing and internationally traded services sectors. Within this set of firms, the Commission will examine six different strata of firms: small and medium sized enterprises (SMEs) in the manufacturing sector which export, SMEs in the manufacturing sector which do not export, large firms in the manufacturing sector, SMEs in the services sector that export, SMEs in the services sector that do not export, and large firms in the services sector. Overall, the Commission expects a response rate of approximately 25 percent. The Commission would like to receive approximately 375 responses in each of the six strata identified above, therefore the Commission will sample 1,500 firms in each of the six strata.

Strata	Total population from Orbis	Total population from Census	Firms sampled*	Expected number of responses
SME exporting manufacturers	14,218	n.a.	1,500	375
SME non-exporting manufacturers	205,873	n.a.	1,500	375
Total SME manufacturers	220,091	169,316	3,000	750
Large firm manufacturers	2,584	4,069	1,500	375
SME exporting services firms	15,191	n.a.	1,500	375
SME non-exporting services firms	610,384	n.a.	1,500	375
Total SME services firms	625,575	562,030	3,000	750
Large services firms	2,780	12,929	1,500	375
<i>Total manufacturing and services firms</i>	<i>851,030</i>	<i>748,344</i>	<i>9,000</i>	<i>2,250</i>
*From Orbis database				

There is not a one-to-one correspondence between the population of firms in the Orbis database, the actual population of firms, or the number of firms captured by Census. Several factors contribute to the differing numbers of firms contained within Orbis and the number of firms reported by Census. Bureau van Dijk, the publisher of Orbis, relies on a combination of publically reported and privately collected data on firms for its database. Although Orbis contains information regarding the ownership structure of firms, in certain cases this information is incomplete, resulting in affiliates being mischaracterized as firms in the database. Orbis therefore overestimates the number of SMEs and underestimates the number of large firms relative to Census, because in certain cases affiliates are not correctly aggregated up to the firm level in the Orbis database. Although the Commission will attempt to target only parent firms in its sampling, the Commission will use question I.3 to screen for any affiliates that may be captured. Additionally, Census classifies U.S. affiliates of foreign parents as firms, while U.S. affiliates of foreign parents are not considered firms in the Orbis database. Finally, Census data are based on

2006 County Business Patterns data, which is the latest available. Bureau van Dijk, on the other hand, updates its database on a continual basis, so the reported number of firms in the Orbis database differs from Census to the extent that firms have entered or exited the market since 2006. The Commission will provide a full description of the data, sampling framework, and necessary caveats in its final public report.

If the Commission finds that certain firms within a given stratum are misclassified and should be included in a different stratum, the Commission will reclassify them, and adjust sampling weights based on Census data where possible. The Commission will explicitly describe all classification and weighting procedures in a statistical appendix of the report.

The questionnaire contains several different types questions:

(1) Proportions, e.g.

- Responses to “Yes-No” questions (e.g. I.9)
- Responses to questions with 4 categories (e.g. I.8)
- Questions on ranks (e.g. II.3), that can be summarized by the percentage of firms listing a response as being “in the top 3”.

(2) Scalar responses, e.g. I.6. (revenue in dollars) and II.3 (Likkert scale I.5)

In both of these cases, the Commission believes that the proposed sample size will yield a degree of accuracy which is sufficient to generate reasonably tight point estimates and confidence intervals. The Commission also believes that the proposed sample size is robust to the possibility of a low response rate.

Survey results in the form of proportions

In the case of proportions, let N be the sample size and Y be the number of observations falling within a certain category, so that $p = y/n$ is the observed proportion in the sample. Then a confidence interval for the true proportion in the population, π , is equal to

$$p \pm z \sigma_p$$

where p is the proportion in the sample, z depends on the level of confidence desired, and σ_p , the standard error of a proportion, is equal to:

Below are presented illustrative 95 percent confidence intervals for various values of the true proportion both for N = 375 and for a “worst case” of a 10 percent response rate (N = 150):

True proportion (π)	N = 375	N = 150
10 percent	6.96 – 13.04	5.2 – 14.8
25 percent	20.62 – 29.38	18.07 – 31.93
50 percent	44.84 – 55.06	42.0 – 58.0

Survey results in the form of scalars

It is a recognized result in the economics literature that the size distribution of firms is approximately log-normal. The relevant literature is reviewed in Luís M. B. Cabral and José Mata, "On the Evolution of the Firm Size Distribution: Facts and Theory," *The American Economic Review* Vol. 93 No. 4 (Sept. 2003), pp. 1075-1090. Distributions of other variables such as export revenues (conditional on revenues being positive) are also likely to be approximately log-normal. Illustrative examples for subsamples of the log-normal size distribution (Cabral and Mata, table 1) suggest an average value of the coefficient of variation of about $\sigma/\mu = 0.7$, with variation in subsamples within an approximate range of (0.4, 1.1).

The most relevant use of scalar data in dollar figures is for differences in group means, for example between exporters and non-exporters or Likkert-scaled questions. The table below indicates the minimum difference between means in various strata for which the null hypothesis of equality can be rejected with different variance structures:

	N = 375	N = 150
$\sigma/\mu = .4$	5.7%	9.1%
$\sigma/\mu = 1.1$	15.8%	25.0%

This indicates both that the degree of accuracy with which differences in group means is enhanced by the proposed sample size, and that the results are still likely to be useful in the event of a low response rate.

2. Collection of information

The Commission will use simple random sampling without replacement to select 1,500 firms from the Orbis database for each of the previously described strata. This stratification is based on the information that the Commission was requested to provide. For example, the USTR asked the Commission to examine the characteristics of firms that provide tradable services including how these firms differ from large exporters of services. As requested, the questionnaire will gather information on trade barriers faced by both goods and services firms. It also gathers information about the global operations of SMEs. The stratification, which does not include all firms, targets firms specific to the scope of the investigation. The stratification will allow comparisons between SME exporters and non-exporters and also between SMEs and large-firms. It will not support inferences about broad categories of any single U.S. industry, such as all automobile manufacturers, as this is not within the scope of the investigation.

The Commission will elicit 1,500 responses from each of the six strata. Given an expected response rate of 25 percent, this will result in 375 responses per strata. Each of the respondents will receive an initial letter from the Chairman of the Commission, explaining the purpose of the data collection, noting that a response by their firm is mandatory, the date by which a response is required, and a Web address where they can access the questionnaire and other materials related to the study including relevant Federal Register notices. A second letter will be sent to non-respondents, several days after the questionnaire is due back to the Commission. Several weeks after the second letter is mailed, the Commission will call firms if response rates are not high enough in specific strata. The Commission is also exploring the possibility of sending postcard reminders to all firms prior to the due date of the questionnaire.

Selection within each stratum will be independent, and each firm within the stratum will have an equal chance of being selected. Given the expected levels of population variances, this number of responses will support calculation of simple statistics, such as the mean, that are within 10 percent of their true value at a confidence level of over 90 percent.

All information in the questionnaire will be used, and different statistics (e.g., means, medians, proportions, etc. for different questions) will be calculated. Cross-tabulations between different questions will be made, and responses between different strata will be compared. In the case of statistics that involve only one stratum, the calculation of the associated confidence level should be straight-forward. In more complex situations, when the responses of several questions are compared, joint confidence intervals may be constructed. Also, different weights may be applied in the analysis to adjust for the fact that the strata are not sampled in the proportion in which they occur in the population.

This survey does not present any unusual problems that require special procedures, and it is a one-time survey.

3. Response rates

The Commission believes that the expected 25 percent response rate is realistic, based on the response to other Commission surveys. As described in points 1 and 2 above, this response rate will provide an adequate level of accuracy and reliability for this study. The survey form itself is short and only requests needed information. With the exception of information about firms' revenue and number of employees, no business proprietary information is requested. It is believed that the relatively simple short form, which only requests a minimum of sensitive information, will help to achieve a response rate of at least 25 percent.

The Orbis database contains information on total revenue and number of employees by firm. The Commission will use this data to determine if there are any systematic differences between respondents and non-respondents. Questionnaires of firms that respond after being contacted a second time will also be compared to those of the initial respondents to ascertain if there are any systematic differences. The Commission will publish the results of this analysis in its report. It is believed that firms will have no incentive to bias their responses as a particular response will not benefit a firm. The Commission believes that these procedures will result in the receipt of a sufficient number of responses of adequate quality.

4. Tests of Procedures or Methods

Once the random samples within each of the strata described above have been completed, the Commission will conduct several diagnostic tests of the sample including comparing means and variances between the sample and the population for each of the strata in order to identify any potential biases or problems with the sampling technique. These diagnostic tests will not require any communication with potential respondents.

Besides the previously described pretesting with firms, the questionnaire has been extensively reviewed within the Commission. Industry analysts and economists who will prepare the report have reviewed the document and made recommendations to improve the content of the questionnaire so that it will include the information needed to adequately address the questions addressed in this study while imposing a minimum burden on the responding businesses. A survey consultant reviewed the questionnaire to assure that it is clear and in a form that is likely to elicit the desired responses.

5. Contact information

Collection and analysis of the data will be the responsibility of the Office of Economics and the Office of Industries within the Commission. William Deese, of the Office of Economics, can be contacted at 202-205-2626. Erland Herfindahl, of the Office of Industries, can be contacted at 202-205-2374. In addition to numerous Commission staff who were consulted on the statistical aspects of the questionnaire, Commission staff also consulted with Heidi Guglielmino, a private survey consultant with Office Remedies. Ms. Guglielmino can be contacted at 703-478-0910.