B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

1. Description of the Activity

The entire population of eligible facilities currently available in the National Industrial Security System (NISS) is used for the voluntary census. The voluntary census is divided into four phrases:

a. Planning: Ensure information such as email address, point of contact (POC) consolidated facilities is correct. Notify industry in advance to allow them time to gather the requirements.

b. Data Collection: Open module via web-based survey tool within NISS. Monitor the census during the deployment period to correct any anomalies before the census is closed.

c. Analysis:

1. Clean the data- inquire with facilities regarding any errant responses, if no response, remove responses.

2. Separate responses and non-responses- The census allows for some facilities to submit consolidated responses on behalf of child facilities. Determine all responding facilities (both parent and child), check against the list of active facilities, and determine all nonresponding facilities.

3. Perform imputation using K-NN algorithm for non-responding facilities. Facilities are categorized using DCSA proprietary classification system. Non-responding facilities will be matched according to category, employee count, and cleared employee count. Closest facilities will be averaged to impute responses for non-responding facilities.

4. Use past census data with actual PSI requirements to project investigation requirements using regression equation linking actuals to census totals. Regression is necessary as multiple facilities may be bidding on same contract while only one facility can win the contract.

d. Monitor and maintain oversight of projected vs actual requirements in order to meet the OMB +/-5% variant; ensure the forecasts remain in-line with PSI budget.

2. Procedures for the Collection of Information

    a. Statistical methodologies for stratification and sample selection;

        NA

    b. Estimation procedures;

        NA

    c. Degree of accuracy needed for the Purpose discussed in the justification;

        NA

    d. Unusual problems requiring specialized sampling procedures; and

        NA

    e. Use of periodic or cyclical data collections to reduce respondent burden.

        NA

3. Maximization of Response Rates, Non-response, and Reliability

In 2019, DCSA transitioned to the National Industrial Security system as the government owned tool to conduct the PSI census to industry. The transition has not been smooth, and the census response rate has decreased due to issues with NISS account access and system latency. Prior to the transition response rates were over 85%, but in the first year of the transition to NISS the response rate was 55%. Since then the response rate increased to 65% 2020, and to 67% in 2021. Ongoing system modifications continue to improve NISS and decrease latency issues.

DCSA, to maximize response rates:

    a. Notifies industry in advance to allow them time to gather the requirements.

    b. Corrects or updates information such as email address, and the point of contact (POC) responsible for filling the census.

    c. Identifies consolidated facilities, a parent facility responsible for its subsidiary facilities.

d. Ensures POC information such as email address for consolidated facilities is correct.

e. Sends e-mail reminders to non-respondents weekly during the data collection period to ensure timely completion of the collection.

To deal with instances of non-response, a K-Nearest Neighborhood (K-NN) methodology is used:

a. Separate non-response from response populations and classify by facility category.

b. Calculate the distance based on the total employee population and total facility clearance population.

c. Calculate average of the requirements from these respondent facilities and assign the values to the non-response facility.

d. Category distribution in the population does not appear to be changing from year-to-year.

DCSA maintains a high working relationship with the respondent field and has consistently experienced high response rates on data collection efforts.

4. <u>Tests of Procedures</u>

DCSA and DoD leadership have approved the PSI-I data collection and analysis methodology to continue to estimate annual PSI-I investigative requirements. The same methodology has been used for over ten years and continues to provide a sufficient estimate of the annual PSI-I investigative requirement.

5. <u>Statistical Consultation and Information Analysis</u>

Primary POC: Chris Pirch 571-305-6241, cell- 540-421-9195 Statistician