SUPPORTING STATEMENT - PART B for

OMB Control Number 0535-0037:

Vegetable Surveys

ICR Authors: Brent Chittenden and Struther Van Horn

USDA, National Agricultural Statistics Service

1400 Independence Avenue S.W.

Washington, D.C. 20250

Supporting Statement B

B.   COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS

1.   Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection method to be used. Data on the number of entities (e.g., establishments, State and local government units, households, or persons) in the universe covered by the collection and in the corresponding sample are to be provided in tabular form for the universe as a whole and for each of the strata in the proposed sample. Indicate expected response rates for the collection as a whole. If the collection has been conducted previously, include the actual response rate achieved during the last collection.

The Vegetable Growers Inquiry sampling frame comprises active operations on NASS's List Frame that have control data for targeted commodities in program states.

The California Processing Tomato Survey sampling frame comprises active entities on NASS's California List Frame that have a processing tomato indicator. The sampling frame is sampled at a 100% rate.

	Vegetable Survey
Survey	Sample Size

2.   Describe the procedures for the collection of information including:
- **statistical methodology for stratification and sample selection,**
- **estimation procedure,**
- **degree of accuracy needed for the purpose described in the justification,**
- **unusual problems requiring specialized sampling procedures**

The California Processing Tomato Survey is sampled at a 100% rate. In contrast, the EOS Vegetable Growers Inquiry is a Multivariate Probability Proportional to Size (MPPS) sampling design[a]. This design takes advantage of the efficiencies of a Probability Proportional to Size (PPS) design, while adding the dimension of utilizing multiple variables in the sample allocation. Additionally, this design makes it easy to target a sample size for every target item.   The target sample size for each commodity was derived using the SRS sample size formula.

For a MPPS design the probability of selection is:

$$\pi_i = max\left(n_1 \times \frac{x_{i1}^r}{\sum_{i=1}^k x_{i1}^{0.75}}, .., n_h \times \frac{x_{ih}^r}{\sum_{i=1}^k x_{ih}^{0.75}} \right)$$

Where:

π_i is the maximum probability of selection for farm operator i.

i is the farm operator,

h is the target commodity,

x is the value of target commodity h, and

r is the power.

The target sample size for each targeted commodity is:

$$n_h = \frac{N_h s_h^2}{\frac{(f_h T_h)^2}{N_h} + s_h^2}$$

Where:

n is the sample size for target commodity h,

N_h is the target commodity population,

s_h is the target commodity standard deviation,

f_h is the target commodity coefficient of variation, and

T_h is the target commodity total.

A sample is chosen by generating a uniform random number (RN), calculating the cumulative probability for unit i as Cumulative$_i$ = Cumulative$_{i+1}$ + π_i, and selecting unit i if Cumulative$_{i+1}$<RN+k< = Cumulative$_i$ for any k= 0,1,2,…n where n is the total number of units selected for the sample[b].

---

[a]Bailey, Jeff and Kott, Phillip (1997)," An Application of Multipurpose List Frame Sampling For Multi-Purpose Surveys, Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 496-500.

[b] Hicks, Susan, Amrhein, John and Kott, Phil (1996), Methods to Meet Target Sample Sizes Under a Multivariate PPS Sampling Strategy, Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 234-238.

Currently, questionnaire content, survey timetables, and survey administration are established by HQ and followed by all States. In addition, there are individual State surveys that are administered directly by the respective Regional Field Offices (RFOs). The master questionnaires submitted are reduced in size to only include the commodities of interest in each State. Data are gathered by telephone interviews, mailed paper questionnaires, faxed questionnaires, personal interviews, and Computer Aided Self-Interview (CASI). Data accuracy and respondent burden are taken into account in conducting the surveys. Because of the variable nature of the vegetable industry, mail lists are frequently updated to ensure maximum coverage.

The sampling frame and sample for the California Tomato Processor Survey surveys are California tomato processing operations. Complete enumeration is employed from the Regional Field Office since there are relatively few processing operations.

At a minimum, State level questionnaires will contain the commodities of interest for US level estimates. In addition, there are agreements in place with State governments that have asked NASS to collect data for vegetables not in the Federal program, so some questionnaires will contain additional crops that do not appear in NASS's national publications. To help reduce burden on both the respondent and on RFO staff, we include these extra crops on existing questionnaires when possible, so we only have to contact the respondent one time each season. The data will be published in State level releases or will appear in the State's Annual Bulletin.

3. **Describe methods to maximize response rates and to deal with issues of non-response. The accuracy and reliability of information collected must be shown to be adequate for intended uses. For collections based on sampling a special justification must be provided for any collection that will not yield "reliable" data that can be generalized to the universe studied.**

    Data reported on vegetable inquiries are used by the National Agricultural Statistics Service to estimate acreage, yield, production, and utilization for crops sold through both fresh markets and to processors. Data are used by many State and Federal agencies and Departments, the vegetable industry, agribusinesses, manufacturing, transportation organizations, educational institutions, foreign governments, and international groups.

    While seeking these data, the Regional Field Offices (RFOs) take great care to keep respondent burden to a minimum. Most contacts have a long-standing relationship with the servicing RFOs. Most fresh market farmers who produce multiple crops are sent one questionnaire listing all of the commodities grown so they only need to complete one form. When more than one questionnaire must be mailed at the same time, they are combined into one mailing.

Information for the vegetable surveys (list frame) is collected via mail, Internet, telephone interview, and face-to-face interview. The mix of data collection modes allows the respondents to reply by whichever mode they are most comfortable with. Large operations and operations requiring special handling (previous survey refusals and inaccessibles, complex operations, etc.) are generally contacted for a face-to-face interview only.

Survey data are subject to non-sampling errors such as omissions and mistakes in reporting and in processing the data. While these errors are not measured directly, they are minimized by carefully reviewing all reported data for consistency and reasonableness. An estimation manual helps in maintaining consistency across surveys and RFO's.

4. **Describe any tests of procedures or methods to be undertaken.**

    The main validation of current procedures is to remain in contact with vegetable growers, processors, and industry leaders and solicit their expert advice. Following each Census of Agriculture (every 5 years), NASS conducts a complete program review of all crop and livestock commodities. NASS verifies which States are included in the different surveys for each commodity in order to produce accurate estimates. Included in this request is an annual fifty interviews for cognitive testing to improve how the respondents provide data.

5. **Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), or other person(s) who will actually collect and/or analyze the information for the agency.**

    Sample sizes for each State are determined by the Sampling, Editing, and Imputation Methodology Branch, Methods Division; Branch Chief is Andrew Dau. Andrew's email is andrew.dau@usda.gov and phone number is (202) 690-8141.

    Survey design and methodology are determined by the Summary, Estimation, and Disclosure Methodology Branch, Methodology Division; Branch Chief is Lindsay Drunasky. Lindsay's email is lindsay.drunasky@usda.gov and phone number is (202) 690-8141.

    The NASS survey administration, data collection, estimation, and publication are carried out by NASS Regional Field Offices; Western and Acting Eastern Field Operation's Director is King Whetstone. King's email is king.whetstone@usda.gov. His phone number is (202) 720-9567. The survey administrators are responsible for coordination of sampling, questionnaires, documentation, training, and data processing.

    Statistics Division, Crops Branch is responsible for national and regional

summaries, analysis, and presentation of data to the Agricultural Statistics Board for final estimates, publication, and the Estimation Manual. The Crops Branch Acting Chief is Anthony Prillaman. Anthony's email is anthony.prillaman@usda.gov and phone number is (202) 706-713-5401. The Environmental, Economics and Demographics Branch Chief is Bryan Combs. Bryan can be contacted by email at bryan.combs@usda.gov or phone at (202) 720-5084.

The NASS census administrator in Headquarters of the Census Planning Branch, Census and Survey Division; Branch Chief is Donald Buysse. Donald's email is donald.buysse@usda.gov and his phone number is (202) 738-3764. The census administrators are responsible for coordination of sampling, questionnaires, data collection, training, Interviewer's Manuals, Census Administration Manuals, data processing, and other Regional Office support.

September 2025