

**UNITED NATIONS
ECONOMIC COMMISSION FOR EUROPE**

CONFERENCE OF EUROPEAN STATISTICIANS

Seminar on New Frontiers for Statistical Data Collection
(Geneva, Switzerland, 31 October-2 November 2012)

Topic (ii): New methods and technologies

**NEW SOLUTIONS, CHALLENGES, AND OPPORTUNITIES: CAPI THE
NASS WAY**

Contributed Paper

Prepared by David Kleweno and Pamela Hird, United States Department of Agriculture (USDA),
National Agricultural Statistics Service (NASS), USA

I. Introduction

1. Budget cuts and decreased resource availability throughout the National Agricultural Statistics Service have created an urgent need to streamline data collection processes and to create a vehicle for remote electronic management of field enumerative staff and data flow. Augmenting the NASS data collection strategy with a Computer Assisted Personal Interview (CAPI) solution offers many opportunities for staff and cost efficiencies, and provides a way to optimize data collection activities. A unique solution was required to meet the challenges of providing security measures to safeguard data, to mitigate risk of broadband signal fluctuation, and to deploy a data entry device that engaged older enumerators who were not technologically astute or computer savvy.
2. The paper describes the development and deployment of a revolutionary CAPI system that leverages private cloud technology, broadband transmission, and use of Apple iPads for data entry without storing data on the iPad hard drive. This innovative solution designed to meet NASS' needs has opened the door to endless cloud-based opportunities to use text, audio, and geospatial input limited only by vision, resource, and budget constraints.

II. History and Research Timeline

3. The National Agricultural Statistics Service is the primary agricultural survey data collection and dissemination arm within the government. Its mission focuses on providing timely, accurate, and useful statistics in service to United States agriculture. Statistical information from approximately 400 reports are provided annually on acreage, production, stocks, prices, income, the environment, conservation, demographics, and other related agricultural sectors. Publications and statistical products must reflect efficiency, effectiveness, integrity, objectivity, impartiality, utility, and confidentiality.
4. NASS utilizes various methods to collect survey data: mail, telephone, personal interview, and self-administered Web instruments. With declining budgets and increasing data collection costs, NASS continues to pursue innovative ways to leverage technology to collect data more

efficiently and effectively. In the 1980s and mid-1990s, NASS Research and Development Division (RDD) explored CAPI as a possible solution. Technology costs and the lack of efficient methods to transmit data in a timely manner made it prohibitive to add CAPI processes to NASS's data collection strategy.

5. In March 2009, RDD once again investigated possible solutions for a CAPI process. RDD made a preliminary recommendation to deploy Lenova netbooks to field enumerators as data entry devices, utilizing broadband air cards for data transmission. The proposal leveraged the use of an "in-house" developed system, NASS Electronic Data Repository (EDR). This system was developed to provide Web access to respondents for self-administered questionnaires. EDR resides in a private cloud environment and is used as a data transmission portal for survey information data flow to and from respondents/enumerators and NASS. Survey questions are electronically selected from another in-house developed system, the Questionnaire Repository System (QRS). QRS assembles the questionnaire and publishes it to the EDR system, rendering a HTML version of the questionnaire to the EDR website. The respondent/enumerator logs into EDR through the Internet and gains access to the questionnaire through a survey code. Once the questionnaire is completed, broadband technology transmits the data to the EDR cloud. Questions are selected from the Questionnaire Repository System (QRS) and published to the EDR system, which renders a HTML version of the questionnaire on the EDR website.
6. NASS launched five broad initiatives in November 2009, to streamline and standardize processes, increase performance, improve quality and create efficiencies across the agency. This multifaceted, reengineering effort was designed to propel the agency forward at a time when budgets and resources were being constrained. In addition to CAPI, the other efficiency initiatives included centralizing network services, centralizing software applications and database processing, centralizing telephoning, frames maintenance and training, and implementing agency-wide video teleconferencing. Implementation of all five efficiency initiatives is required to maximize benefits, although centralized telephoning and CAPI are the main drivers.
7. The CAPI project manager was named and a team was formed in November 2009. The project mission was to prove the concept for the CAPI initiative and, if feasible, implement CAPI in all states by October 2014. The objective was to implement in a cost neutral manner with an expected return on investment (ROI) of 3-5 years.
8. A number of organizations are successfully using CAPI data collection. The most widely used or standard CAPI method of data collection involves an interviewer going to a respondent with a questionnaire loaded on a computer. Survey questions are asked with responses stored on the device. The information is then electronically transmitted to a data collection site. Major benefits include improved timeliness of the data, increased data accuracy, reduced data entry requirements, flexibility in making enumerator assignments, printing limited to only back up questionnaires, and greatly reduced postage costs needed for mailing completed questionnaires to a data collection center. Implementing a CAPI solution at NASS will greatly increase the efficiency of data collection activities when leveraged with concurrent initiatives to centralize systems, processes, and data collection/handling activities.
9. Adding the CAPI method of data collection offers many opportunities for staff and data efficiencies as well as the benefit of an electronic field data collection program. However, NASS had to develop a unique solution to address three key challenges: having the necessary security measures to safeguard data, mitigating the risk of broadband signal fluctuation, and meeting the needs of its enumerators, who are largely retirees with an agriculture background and typically not computer savvy.

III. Pilot to Implementation

10. The CAPI project was designed to move through four phases: Proof of Concept –Fiscal Year (FY) 2010, Pilot – FY 2011, Implementation – FYs 2012-2013, and Closeout – FY 2013.

- (a) In FY 2010, a risk and opportunity analysis and business case was developed. Testing, training, and data collection in two states using netbooks and aircards determined that the concept worked but the tools were less than optimum. NASS researched, tested and then incorporated Apple iPads for data entry and personal hotspot WiFi devices for data transmission. This created a user friendly, minimal-learning-curve solution. Ajax technology was incorporated to mitigate signal fluctuation and loss. This solution proved effective but required rewriting the EDR system to incorporate Ajax technology. The proof of concept phase was completed in FY 2010 with approval given to conduct a pilot in FY 2011.
- (b) The primary goal of the FY 2011 pilot phase was to complete testing of major survey instruments in the CAPI/EDR system and expand the scope to include eight field offices. A major efficiency was gained by redesigning the training program from 24 to 12 hours and maximizing the use of video teleconferencing. The training workshops integrated the ROPES¹ Lesson Plan Model, which focused on five training steps—Relate, Overview, Present, Exercise, and Summarize. This method helped engage the NASS enumerators. WiFi devices were replaced with new iPad2 Apple technology, which incorporated a built-in 3G device. Cost and time requirements met projections and goals.
- (c) A significant step was taken in FY 2012 with the goal of implementing CAPI in 46 states by December 2012, one year ahead of schedule. Training programs were refined again to adjust for the New Apple iPad and Apple Operating System (AOS) as well as 4G technology. Advancements were made to incorporate centralized processes associated with the agency initiatives. Survey data collection expanded to more than 1,700 iPads encompassing over 3,000 survey questionnaires. The initial implementation hardware costs were offset by reduced costs for postage, printing, and data entry time. A project management Project Evaluation and Review Technique (PERT) analysis showed return on investment projections when fully implemented of more than \$3 million annually, based on staff hour savings, not including printing or postage. Project completion goals were on target.
- (d) Completion of the implementation phase will be followed by a closeout evaluation of the initiative. This assessment will include establishing the future operational framework for the CAPI program. As the project evolved, numerous changes have been made in the technical and resource requirements. Policies for training, equipment, security, quality, and cost assessment factors need to be developed and implemented as the program continues to develop and mature. Cost savings measures, data collection procedures, and metrics for success will be enhanced and refined. Finally, a Lessons Learned document coupled with recommendations for improvement will be written. FY 2013 will usher in the opportunity to further streamline, enhance, and develop CAPI processes.

IV. Challenges and Solutions

A. Initial Challenges

- 11. Research and testing had established a valuable platform to re-evaluate and initiate the CAPI project. Technology advances made it feasible to initiate the project but not without numerous challenges.
 - (a) *Technology* – The original RDD recommendation based on Lenova netbooks and aircards was less than optimum. End-users (enumerators) tested the netbooks and determined they were too heavy to hold for long periods, battery life was limited to 4 hours, enumerators experienced acute difficulty in manoeuvring through the survey instruments using a mouse pad, and the netbook graphics were not sufficient for use outdoors in sunlight. In addition, the aircards proposed for data transmission overheated, did not hold a signal for long periods, and, because they protruded from the side of the netbook, were easily broken.

¹ROPES Lesson Plan Model – R. Carkoff/Moshinskie

- (b) Fluctuation of signal for broadband transmission – Signal stability during an 8-hour work day was a major issue. Research and on-the-job testing showed the data signal was present at some point during the day but was inconsistent. A variety of issues caused signal fluctuation including, trees, topography, power lines, and location of the interview site.
- (c) *The original EDR system* – This software was developed for respondents to use from a home computer and was designed to work across all platforms and methods of data transmission. Downloading questionnaires proved to be extremely slow. Enumerators found this to be a significant challenge as multiple questionnaires needed to be downloaded in a short amount of time.
- (d) *Security* – A major requirement was to meet all security concerns both at NASS and at the Department of Agriculture level. Federal mandates had to be addressed. NASS was the first federal agency to put iPads into production and one of the first federal agencies to adopt private cloud-type technology. This meant testing and developing procedures which were undefined. The final solution had to address all aspects of security and ensure the data were protected during collection and transmission. The CAPI/EDR system could not be located behind a secure firewall since it had to be public facing for access by enumerators who are not federal employees.
- (e) *Budget* – The project was slated to be cost neutral, which restricted the funds available to spend on the devices and later the implementation plan. Budget constraints drove several key decisions as outlined below in the Initial Solutions section, paragraph 12 (e).
- (f) *Resistance to change* – CAPI was not a new concept to NASS. Since the 1980s, NASS had researched and tested several options. The difficulties encountered in previous efforts tended to bias against acceptance of any CAPI solution. Personalities and business culture at NASS resisted altering a business process that had worked for a hundred years. Change required moving away from well-known and defined processes—and adding CAPI to the data collection toolset promised radical changes. Resistance from the enumerators was based on fear of the unknown and apprehension regarding their ability to perform the necessary functions with a computerized device. The typical demographic of an enumerator is a retiree with an agriculture background and limited experience with information technology. Lastly, there was concern whether respondents would accept the computerized data entry method.
- (g) *Training program* – The training program was a key component to the successful implementation of CAPI. The program needed to be flexible and able to address changes when they occurred. The program also had to be engaging and incorporate training methods appealing to adults. In addition, budget constraints and restricted resources put the training program on a very tight time schedule and limited the number of attendees. Finally, the training program had to be standardized to address quality concerns.

B. Initial Solutions

- 12. Sound business management principles were used to address major programmatic problems. Innovation, vision, and determination ultimately produced viable and functional solutions.
 - (a) *Business Case* – A formal document was created and presented to senior management to address NASS's unique requirements for the data entry and data transmission devices. Extensive research and testing of devices both within the agency and with field enumerators was conducted over several months. Devices tested included convertible netbooks, laptops, and tablet type devices but the Apple iPad was the only device that met all technical, ease-of-use, and cost requirements. Personal hotspots from Sprint and Verizon were selected for data transmission. From the beginning, Apple, the broadband providers, and the equipment vendor, worked closely with NASS to overcome incompatibility problems and other issues that surfaced throughout implementation. NASS incorporated new iterations of iPads as they evolved, including the current solution with 3G/4G built into the device; strong business relationships helped to ensure

the success of each technical upgrade. Flexibility was built into the solution so that each part of the CAPI solution could be easily replaced to incorporate new and improved technology or processes. The CAPI team continues to research and test emerging technologies looking for additional savings and innovative solutions for continuous improvement to the CAPI process.

- (b) *Fluctuation in the broadband transmission signal* – Mitigating signal problems required a two-step solution. First, enumerators have to download survey questionnaires into an open browser window on the iPad when a strong data signal is present. This download normally takes place at the enumerator’s residence or at a location during the day with a strong signal. Second, to facilitate the transmission of data, Ajax technology was attached to every item cell box where data is entered. Ajax constantly sends out a transmission request to locate a data signal. Once the data is entered, Ajax searches for a signal and immediately transmits the data when a signal is found. The transmission only requires a few seconds. The open browser window is closed once the data for the entire questionnaire has been transmitted by Ajax. This solution is significant in that it allows NASS to conduct surveys as though on the Internet whether a data signal is present or not. The use of browser windows and Ajax enables the enumerator to go to the next interview and enter data whether previous survey data has been transmitted or not. As needed, the operational memory or RAM of the iPad is used but without storing data on the hard drive of the device.
- (c) *The original EDR system* – Slowness in downloading questionnaires on the EDR system was resolved through caching the questionnaires in memory, coupled with increased transmission speed using the personal hot spots, and later 3G/4G, and incorporating Ajax into the code. To take full advantage of these technologies, NASS optimized and rewrote much of the EDR system code.
- (d) *Security* – NASS, which did not have the resources to facilitate and manage security passes (HSPD 12 Linc passes) for over 3,000 non-federal enumerators, had some unique security concerns. An innovative solution was needed that did not require admittance to systems behind the NASS firewall or require NASS to store information on the devices’ hard drive. The solution was to leverage cloud-type technology to use a private cloud with public access. A number of measures were put in place to further address security concerns. First, secure socket layer protection was incorporated into the EDR site. Extensive testing by an outside firm revealed the EDR cloud could not be penetrated and garnered an AA rating for penetration tests. Secondly, NASS modified the iPad settings to introduce a complicated password, device lockdown for application downloads and cookie installation. Lock down processes were also put in place for several non-use icons that store information on the hard drive, such as the mail icon, Facetime, Notes, and the calendar. As an additional measure, NASS developed a system to test userids, passwords, and user keys for authenticity in order to access any data in the EDR cloud. Four levels of security are validated prior to any download of survey questionnaires or any access to data. Finally, every user was required to take security training and a test to reinforce security measures, and sign paperwork regarding proper usage of the iPad before receiving an iPad.
- (e) *Budget mitigation measures* – Budget constraints played a major role in several key decisions. The first decision was to implement a data entry device with a budget level of \$1,000 or less. The training program was refined and streamlined from 24 hours to 12 hours with implementation on a regional basis to save training and travel dollars. The last key decision was to forgo replacement insurance on the iPads. Using project management analysis, it was determined the projected replacement cost of 25 percent was more cost effective than an annual insurance cost of \$100 per unit.
- (f) *Resistance to change* – This initial challenge was both expected and difficult to address. The solution was to adopt a strong marketing plan to gain “buy in” at all levels. Numerous presentations focusing on benefits were given at workshops for enumerators and management workshops for NASS personnel. The internal NASS “newspaper” published several articles outlining the benefits and efficiencies to be gained from the CAPI project. Successes and ongoing plans regarding the project were communicated on a regular basis. Enthusiasm was generated at

the department level through blogs, webinars and published magazine articles. Positive feedback and endorsement by senior management as well as the department was shared with all NASS staff and the enumerators. Extensive testing was done in the field with office staff and enumerators to address their concerns and to ensure that the solution would address their needs. They were given credit for participating and their suggestions and feedback were implemented into the solution. A general feeling of ownership was cultivated among the enumerators, field office staff, and CAPI team. Even the vendors were treated as part of the team. This feeling of ownership contributed widely to the initiative's success. Generally, those who came in contact with the CAPI solution had a genuine desire to see it succeed. Overall response from respondents was enthusiastic and very positive. Minimal effort for "buy in" from respondents was required.

- (g) *Training program* – CAPI team members travelled to each field office to conduct workshops for enumerators in the state. These members trained a core group of enumerators and 2-3 field office staff on CAPI procedures through a standardized training program. The workshop incorporated the ROPES¹ Lesson Plan Model to maximize training comprehension and understanding. The workshop consisted of a variety of training styles including presentations, storytelling, hands on practice exercises, and mock interviews. Once the training was completed, the field office was responsible for training remaining enumerators and any future enumerators on CAPI, utilizing the core group of enumerators as resources. Training materials were made available to each field office, although standardizing the materials proved somewhat problematic as they needed to change as technology and enhancement changes were incorporated within the CAPI solution. This train the trainer or transference approach did reduce the quality and consistency of the training program somewhat. The preferred method to ensure quality and consistency was to have one or two teams train all enumerators, but this was not feasible, due to budget constraints and time limitations. As conducted, the training program was a compromise of the standard train-the-trainer approach. It provided quality and consistency to a large part of the resource base in each field location. The training program was extensively tested and streamlined through trial and error and feedback from enumerators and pilot states over an 18-month period. The original three-day workshop was cut in half due to budget limitations. The current training program incorporated most known methods of learning and was incredibly effective considering the time constraint of 12 hours. Feedback from each workshop indicated enumerators felt rushed and did not have enough practice time. To address this need, a separate training site was created enabling enumerators to access surveys for practice at all times. Enumerators were given a 4-6 month transition time to move from conducting survey interviews using paper questionnaires to conducting the interviews using iPads. This policy was communicated in training workshops and reflected in the training manual. Feedback from field offices and questions reflected a need for improved knowledge of the EDR system and also some legacy system components. The CAPI team developed a video teleconferencing training program for the field offices to address these issues. Overall training feedback was extremely positive from field offices and enumerators.

C. Current Challenges

13. As with all change, challenges continue. Some challenges result from solutions implemented to mitigate initial challenges; some are issues that continue throughout the project life cycle. The CAPI team continues to maintain fluidity and flexibility to address new issues and opportunities as they arise.
- (a) *Budget* – Budget constraints are a constant. NASS and other federal agencies strive for innovation while maintaining sufficient funds for operations. The team continues to monitor technology for cost efficiencies and time-saving features.
- (b) *Technology* – NASS continues to replace older iPads with new iterations or models. This policy has led to some unique challenges involving incompatibility issues with broadband providers and a lack of ready inventory for purchasing. The accelerated project timeline to implement all states before the Census of Agriculture constrains testing of new technology and puts NASS in a reactive position instead of a proactive position. Testing new development enhancements and working closely with vendors are required to address incompatibility and inventory concerns.

- (c) *Training program* – The training program continues to require streamlining and updating to address new technology upgrades and enhancements to the EDR system. As new processes are added, updated training and information must be provided to previously implemented states. Documentation memos and publication to the NASS Intranet are the main methods for communicating updates. The team continues to address various inconsistencies that arise from training since workshops were conducted with multiple training teams, which affected quality and consistency. Even though training materials were standardized and train-the-trainer techniques were used, some differences continue to surface based on personalities and training techniques of the training team.
- (d) *Limited resources* – The CAPI team consists of four full-time staff positions and five part-time positions, supplemented as needed by specialized NASS personnel. The accelerated timeline created some challenges for conducting workshops with the limited staff resources available. At times development enhancements were postponed and on occasion management resources were used as trainers to help support the workshops. Field office staff also provided assistance during hands on exercises. In addition, competing priorities for the same development resources continue to be an issue as necessary enhancements and “fixes” to CAPI survey instruments and the CAPI/EDR system are sometimes delayed due to resource constraints.
- (e) *Integration of data with legacy systems* – Many of the legacy systems at NASS are in the process of being updated or centralized. Once all the new systems are in production, there will be a seamless transmission of data from CAPI processes to other systems used for survey management, editing, analysis, and summarization. Many systems in development and/or testing require the same resources as the CAPI project. Since CAPI is in a production environment, there is a time lapse where the old legacy systems are still in use. The CAPI team has been innovative in creating workarounds to integrate the data from EDR with the legacy systems. However, the development of these workarounds compromises resource availability to work on future development and enhancements for the CAPI solution. In time, the project timelines will mesh and this particular issue will be resolved.
- (f) *Project deployment* – The project was deployed in phases using Lean Six Sigma project management principles. This concept allowed for continuous improvement of the CAPI process and for testing the survey instruments in the CAPI/EDR system through data collection activities. Each phase built upon the successes and feedback of the previous phase. Testing more than 3,000 NASS survey instruments in the CAPI/EDR system prior to implementation would have delayed the project indefinitely. An innovative approach of using the field enumerators and field offices to test survey instruments allowed CAPI implementation to move forward and to accelerate the project plan. However, this method is not without drawbacks. The concept met some resistance as NASS historically tests and validates processes before putting them into production. Time is proving the effectiveness of this approach combined with training that reinforces testing of the instruments. A full-time staff position was established to help mitigate the testing issue for the major functions of the survey instruments. Attempting to address both implementation and production issues as they surface remains a challenge.
- (g) *Infrastructure instability* – Implementing technology for the CAPI solution continues to raise novel problems. There are no established benchmark tests for server functionality/capability that address the innovative solution put in place. Consequently, as more users are added, server issues continue to increase exponentially, creating an unstable server environment. A load balancer and additional memory were added to the core solution to address connection and data transmission requests. However, configuration errors, server hardware malfunctions, and issues with timeouts and slowness continue to cloud the CAPI process. This instability is compromising CAPI staff resources that ingeniously create and implement workarounds for production server issues. New servers are being installed to increase the CPU capability by 12 times the current solution. Load balancing is being distributed over five servers instead of two. Projections through robotic tests show this solution should resolve all issues with the current portfolio of available surveys but the

addition of more complex surveys and additional users may require additional hardware improvements.

- (h) *Scope creep* – The enthusiasm and successes of the CAPI team and project are two sided. The demand for additional CAPI features continues to increase as the project moves through implementation. The team’s creativity, innovation, customer service orientation, and willingness to address suggestions and improvements led to developing numerous features and workarounds to facilitate data flow and provide automated processes for cost and time efficiencies that were outside the scope of the original project. In addition, new projects are under way that must interface with the CAPI structure and require the expertise of the CAPI team. Although exciting and of great value to NASS, these new projects and demands strain resources that are already severely limited due to the accelerated project plan. This amplifies the challenges the project is experiencing and the constraint of sharing resources with other efficiency initiatives at NASS.
- (i) *Assimilation into the organizational structure* – The inclusion of CAPI procedures into the data collection toolset has revolutionized data collection at NASS. Policy changes to address survey coordination, respondent burden, and survey procedural changes are required to fully realize the efficiencies to be gained. These policy changes take time and require collaboration with other organizational units. In addition, other groups are exploring new uses that will be delivered and coordinated through the CAPI vehicle. Policy regarding testing, production, and deployment have not been finalized, which increases risk with regard to security, standardization of training, and miscommunication of applications available to enumerators on the iPad. These issues and concerns continue to evolve and develop over time as CAPI processes become standardized and more opportunities for efficiencies are discovered. A recommendation to create a new organizational structure for CAPI has been proposed to senior management. The project reaches completion once implementation and closeout documentation occurs. The CAPI processes will then transition into a CAPI program for future development and expansion that includes large, complex surveys not currently available, policies and standards, and exploration of opportunities for new efficiencies.

I. Program Efficiencies and Opportunities

A. CAPI Outcomes and Benefits

- 14. The improvement and gains seen in data collection methods and survey processes have been significant during the developmental and early implementation stages of CAPI. As the project matures and technology advances, greater cost savings and efficiencies will be realized. In these early phases of CAPI development and implementation, NASS realized benefits in eight areas.
 - (a) *Time* – This variable is at the center of every survey process, often restricting the survey administrator’s ability to change or maximize benefits and control cost variables during the data collection process. The CAPI vehicle provides a structure to monitor data collection variables and introduce appropriate changes for positive outcomes since data are available for review within hours. Field enumerators enter respondent data at the point of interview. If the interviewer chooses not to do this due to respondent or survey constraints there is still flexibility to input data off site on the same day. If needed, paper can remain the fall back in special situations. Access to live data allows for flexibility in managing survey workloads, immediate evaluation and review of data, opportunity for data verification, and if needed follow-up and ultimately an increase in response rates and accurate estimates.
 - (b) *Cost* – Budget constraints and reduced resource availability served as the NASS initiative for implementing CAPI processes. Streamlining processes and lowering data collection expenses were tied directly to the CAPI initiative with an aggressive goal of returning a positive return on investment within three to four years of full implementation. Full implementation will be achieved 6 months early. Immediate printing and postage savings were seen during the pilot and implementation phases as additional surveys were administered under CAPI and the number of states expanded. A sizable savings was also achieved by eliminating the need for external

contracted data entry support. Projected benefits were also realized in staff time savings associated with pre-survey administrative functions, survey check-in, paper editing, data keying, and post survey tasks. A PERT (Project Management Model) completed in 2011, which evaluated only time savings associated with 12 major surveys, projected a cost savings of approximately \$3.2 million annually, once all states were implemented. The analysis results were conservative as it did not reflect paper, printing, postage, and survey savings for surveys beyond the 12 evaluated as well as large enumerated surveys that will be included in the future.

- (c) *Quality* – Opportunities to monitor, evaluate, and establish quality assurance parameters and matrices using CAPI have been initiated. As will be explained later, CAPI opens a tremendous resource avenue, waiting to be developed and mined. Preliminary development assessments have reflected that the use of iPads for immediate data entry results in fewer errors and more consistent data. A minimal number of edit checks have been used to date as the primary emphasis has been on expanding the number of states using CAPI. An increase in response rates has been seen, due in part to the extended number of days available for enumeration. Interestingly, the novelty of this technology has intrigued respondents in such a way they have shown an openness to participate in the survey, while previously showing resistance. Development of quality assurance measures involving CAPI opens a rich field of exploration with current results reflecting many viable and positive outcomes. The CAPI project incorporated both project management and Lean Six Sigma principles.
- (d) *Communication* – The success or failure of programs, projects, and survey data collection strategies depends on immediate access and availability of accurate communication. CAPI serves as an excellent vehicle for enhancing and broadening this connectivity. The CAPI solution functions as a communication hub between field enumerators, supervisors, regional offices, data collection centers, and when needed headquarters. No longer is there a need to wait for data review or evaluation. Access to respondent information can be obtained as soon as the interview is complete. Respondents will be able to see the latest reports on the NASS website during the interview. Training materials are directly accessible by everyone in the communication path. Connectivity extends to administrative functions as well. Most communication boundaries and limitations have been removed with the flexibility, accessibility, and timely flow of information now available using the CAPI structure.
- (e) *Administration* – Electronic access to administrative and respondent data provides management with the flexibility and opportunity to monitor and control survey processes as they occur. Field assignments are electronically sent to supervisors and enumerators. Along with field office staff, the national data center, and headquarters, supervisors can adjust and shift assignments and monitor enumerator progress electronically. Benchmark modifications can be made as the survey progresses relative to time, quality, and performance. Field work can readily be shifted to different data collection modes. Administrative time, mileage, and related expenses are electronically recorded by the field enumerator using a private Web site. No information is stored on the iPad. Immediate access to both the administrative and survey data allows for real time and quality metrics assessment. Management has the opportunity and flexibility to make strategic data collection and survey changes as needed to meet cost, efficiency, and quality goals.
- (f) *Program Initiatives* –The CAPI program is maximized by the effective implementation of the five initiatives established in November 2009.
- Each of these program enhancements blend with plans to restructure NASS’s headquarters and 46 field offices. Consolidation of data collection, training, list maintenance, and survey processing were moved to a National Operations Center. The objective was to streamline survey and estimation functions and provide consistent, repeatable, quality-based outcomes. Success of the restructuring and downsizing effort depended extensively on the successful completion and integration of the five initiatives. CAPI was a major contributor to the streamlining effort and quality enhancement undertaking. Cost and efficiency gain projections were directly related to the successful regionalization of staff and data resources.
 - CAPI was also an integral factor in establishing a structure that reaps the benefits of a centralized data collection and processing system. Flexibility and functionality of the iPad

served as the catalyst for developing a solution that successfully marries current technology with the NASS operational program.

- (g) *Data Flow Efficiency* – A fully functioning centralized environment, the result of the five initiatives, provides the venue for maximizing the benefits of an efficient survey processing data flow. The centralized survey management system incorporates numerous applications/databases across NASS to integrate survey specifications, sampling, data collection, processing, and output systems. CAPI links to these functions providing the required technology and connectivity to deliver survey data in a timely, cost efficient, and quality controlled manner.
- (h) *Research Opportunities – Quality and Time Efficiencies*. CAPI-related research plans have been established to enhance and strengthen the NASS data collection program. Timing has been contingent on implementation of CAPI in all states. Initial research efforts will focus on broadening the functionality of the iPad and exploring ways to strengthen statistical survey data collection protocols related to response rates and quality assurance.

B. Future Components and Vision

- 15. The scope and capacity for using CAPI to strengthen, manage, and develop the survey framework continues to expand. With CAPI fully implemented, NASS will focus resources in FY 2013 on opportunities to increase its scope to large complex surveys while exploring and developing components that integrate survey efficiency and quality.
 - (a) *Complex Enumerative Surveys* – A significant gain in survey effectiveness and efficiencies, with the inclusion of broad-based data input edits and cost control, will be realized with the inclusion of complex surveys. In FY2013 complex enumerative surveys will be developed that extend to more than 40 pages, involve tables, and cross reference items. The implementation phase focused on simple surveys with a relatively straightforward question design, taking less than an hour to complete. NASS conducts several of these large surveys involving extensive coverage of production, environmental, chemical use, and economic assessment factors. The Census of Agriculture will also be available for field enumeration after other avenues, such as mail and telephone response, have been exhausted. Significant cost savings will be generated from reduced paper, printing, editing and data entry time.
 - (b) *Objective Measurement and Yield Surveys* – Specialized survey work that involves conducting field observations of crop progress and production are performed for several commodities both for national and state-specific programs. Paperless entry of on-site observed measurements and counts will be recorded and immediately transmitted to the regional field office. Prototype testing for Valencia oranges is under way in California.
 - (c) *Geospatial* – Linking sample information with related maps or geo-reference points will open the door to several survey efficiency opportunities. Survey specifications will require locating and identifying field samples on a map, drawing boundaries, and collecting data associated with the sample. An application wrapper for Safari will be developed that allows enumerators to move seamlessly between a questionnaire and map, using tabs on the iPad. No map information will be downloaded on the hard drive, thus adhering to security guidelines.
 - (d) *Survey Management* – Once enumerated surveys are under way, administrators have typically had limited flexibility in making changes to the course of the survey. Current, up-to-the-minute knowledge is lacking relative to survey progress, sample completions, assignment adjustments, and most notably costs. CAPI removes these limitations. No longer will survey administrators have to rely strictly on benchmark and interim reports provided periodically during the survey cycle. Everyone along the data collection chain will have consistent and current information on the full status of the survey. Supervisors will know exactly what has been completed and be able to validate survey reports with their staff via the EDR cloud with access through the iPad. Quality reviews of reported data, survey workload adjustments, and cost analysis can be done in real time. Administrators in the regional offices and national operations center will have the same

information, at the composite level or, if desired, by enumerator. The CAPI solution is a powerful tool providing comprehensive centralized survey management tracking and control functionality.

- (e) *Time & Attendance (T&A)* – Enumerators will use the iPad to record time, mileage, expense, and survey sample related information at the conclusion of each day. Information will be edited when entered, reducing errors and secondary key entry time. Cost and expenditure matrices will be in place to monitor survey performance. When the program is mature, survey administrators will be positioned to evaluate survey costs each day. Flexibility will exist to adjust and adapt programs proactively to changing budgets, response rates, survey design, and other factors.
- (f) *Training Resources* – CAPI will greatly expand and enhance training opportunities and communication with field staff. Materials will be available online virtually eliminating paper documents. As program changes develop or need to be updated, they will be pushed immediately to all office and field staff. At some point, a question tab may be associated with each survey cell within a questionnaire, allowing field staff to review related instructions and background information during an interview, thus positively affecting data quality.
- (g) *Public Relations* – Visual or sound bite material will be easily accessible using the CAPI structure to share directly with the target audience. Information will be tailored to the audience to have the greatest impact. Whether at a trade show, interview, or event, current point of contact information will be available. Improved response rates are expected as respondents recognize the survey benefits.
- (h) *Data Quality, Paradata, and Editing* –
 - CARI – Computer Audio Recorded Interviews (CARI) saves sound files and screenshots for a sample of questions to be evaluated for quality assurance purposes. The iPad recording capability serves as a potential tool in the future to do interviewer evaluations, data quality assessments, evaluate question wording and placement, and aid as a training tool.
 - Adaptive Design – NASS is using paradata in probability county estimates data collection procedures. Responsive design methods are being established to (i) pre-identify a set of design features potentially affecting costs and errors of survey estimates, (ii) identify a set of indicators of the cost and error properties of those features and monitor those indicators in initial phases of data collection, (iii) alter the features of the survey in subsequent phases based on cost–error trade-off decision rules, and (iv) combine data from the separate design phases into a single estimator.
 - Propensity Scoring – NASS is developing a system to use paradata for managing field interviews. Nonresponse propensity scores will be used to identify highly likely nonrespondents before data collection begins. This will allow NASS to alter data collection procedures, target publicity materials and incentives, and adjust interviewer assignments to focus on nonresponse avoidance.
 - Editing – CAPI provides an expanded opportunity to implement an edit repository system. Each cell would have the option of having associated edit parameters. A step further would consider the possible use of previously reported information, administrative data, and ancillary records in survey interviewing. Edits, skip patterns, and use of previously reported data will be incorporated as the CAPI survey process matures. The objective of increased accuracy and reduced respondent burden must be weighed against the potential for systematic reporting errors.
- (i) *Instantaneous Communication* – Protocols and processes will be established that allow for direct communication among field staff, supervisors, and administrative staff. Security guidelines will need to be addressed to permit communication methods such as instant messaging and emailing.

C. Conclusion

16. Implementation of this innovative CAPI solution totally revolutionizes the NASS enumerative survey data collection framework. The cloud and iPad technologies streamline processes, reduce

staff and ancillary resource requirements, and unlock an extensive array of visionary opportunities. Enhancements and advances are staged for future development, all within the scope of current cloud technology and agency security requirements. Collecting and transmitting data through the EDR via private cloud technology provides significant cost savings in printing, postage, and staff requirements sufficient to show a 3-4 year return on investment. Increased data quality, improved response rates, and more timely survey data are precursors to additional benefits of including complex surveys, electronic access to training and public relations material, survey administrative management, and data quality assessments. CAPI, leveraged with concurrent NASS initiatives to centralize systems and processes, provides endless cloud-based opportunities to use text, audio, and geospatial input. CAPI provides a functional infrastructure designed to flexibly adapt to future technology and NASS's business needs.