

Setting a Research Agenda

Linda J. Young
Department of Statistics
University of Florida
Gainesville FL USA
LJYoung@ufl.edu

Dale Atkinson
Research and Development Division
National Agricultural Statistics Service
United States Department of Agriculture
Fairfax, Virginia USA
Dale_Atkinson@nass.usda.gov

Mark Harris
Research and Development Division
National Agricultural Statistics Service
United States Department of Agriculture
Fairfax, Virginia USA
Mark_Harris@nass.usda.gov

ABSTRACT

Developing a research agenda in a government agency with numerous ongoing programs is challenging. Operational units make requests as needs arise, and emergencies can disrupt a carefully prioritized research plan. Individuals or teams can be tasked with meeting the research needs. Matching people with the requisite skill set to a specific project requires an understanding of personnel. In this paper, setting and implementing a research agenda within a government agency is discussed in the context of the U.S. Department of Agriculture's National Agricultural Statistics Service.

Keywords: Surveys, Statistical research

1. Establishing Research Priorities

The process of identifying research topics and developing a research agenda is a challenging one. Within the Research and Development Division (RDD) of the National Agricultural Statistics Service (NASS) in the United States Department of Agriculture (USDA), this process is complicated by the fact that RDD's research efforts are a combination of 1) short term analysis requested by NASS operating units, 2) short to medium term reimbursable statistical/survey consulting for other USDA agencies, and 3) medium to longer term efforts to address current concerns in support of the agency's operational efforts, 4) longer term research, geared toward future improvement in the agency's survey/estimation processes. The requests to RDD originating from operational units come on an on-going basis, further complicating the prioritization process. RDD's management makes strong efforts to stay connected with and engaged in the operational issues to enable it to stay abreast of the new and emerging issues. External reviews of NASS

programs identify additional areas for further research. As examples, the Agricultural Research Management Survey program (National Research Council (2007)), the Agricultural Prices program (The Council on Food, Agricultural and Resource Economics (2009)), and the Census of Agriculture program (The Council on Food, Agricultural and Resource Economics (2007)) have each been reviewed by external experts within the past four years. Based on those reviews, NASS has increased its efforts toward direct research and process improvement. The top priorities identified for research during the fiscal year that just ended on September 30th are as follows:

- 1) Number of farms estimation research that helps explain and eliminate (or reduce) biases that appear to be present in the agency's survey program.
- 2) National/state level modeling, with estimated error bounds, in support of the Agricultural Statistics Board's estimation process for key acreage/production and livestock characteristics.
- 3) Small area (i.e., county) estimation of crop yield/production and cash rents – characteristics crucial to farm program administration.
- 4) Use of statistical/selective editing in the agency's sample surveys to improve both efficiency and data quality.
- 5) Continued expansion of the agency's remote sensing program in support of its crop acreage and yield estimation programs and to provide additional products of great interest to both the agricultural community and the general public.

Periodically RDD has formally canvassed the agency as a whole to identify topics of research interest/need, prioritized these, formed a research agenda, vetted it through its Strategic Planning Council of agency senior management, and distributed it throughout the agency. Although there is benefit to doing this occasionally, it is a time consuming and somewhat ungainly process that ultimately results in a research agenda containing a wide array of projects, some of which will already have been accomplished at the time the research agenda is finally published. Including these finished projects in the final publication is awkward, but omitting them results in a product that seriously under-represents the work that the RDD actually does. There is also the other side of this – often major agency concerns that must be addressed through research emerge soon after the agenda is “set.” These must then be retro-fitted into the agenda, resulting in a process that is never completely set, but constantly evolving.

To facilitate the communication of the scope and status of its research efforts throughout the agency, RDD's current five-year research plan has been populated into an interactive “project tracker” utility that resides on the agency's Intranet. On-going projects are updated, new ones are added, and completed ones are “retired,” on a quarterly basis. This is definitely beneficial in providing transparency to the division's research efforts. However, the central problem remains that the combination of the required mixture of short and longer term research efforts and the continual emergence of new, important research issues that need to be addressed make “finalizing” a research agenda the challenging work in progress that it continues to be.

2. Matching Personnel to Research Topics

Once the areas for research emphasis have been determined, RDD has to determine how best to address them. Often the required personnel are available, and it is simply a matter of directing their effort. Although several promising statisticians work within RDD, a number of the personnel are young, lack research experience, and need mentoring so that they can assume more primary responsibilities. So, research efforts are often conducted in teams and the more experienced individuals are given responsibilities for directing them.

NASS's Research and Development Division is housed in Fairfax, Virginia, within close proximity to Washington, D.C. This provides the opportunity to bring in area statisticians to help with projects. Partha Lahiri, a statistics faculty member at the University of Maryland and the Joint Program in Survey Methodology, has spent the past two years working on small area estimation so that estimates of crop acreage and yield at the state and local (county) levels are more precise. He has worked directly with several RDD staff members and has presented interim results to both RDD and operational staff in the agency. Other statisticians may be brought from more distant places if there is a good match between their expertise and agency needs. Malay Ghosh, a statistics faculty member at the University of Florida, has recently begun working on county estimates of cash rents, which again focuses on county modeling estimation, but with different inputs and constraints. A new Ph.D. has been recruited by NASS, and a portion of Malay's responsibility is to help mentor her.

Finally, NASS has developed a cooperative agreement with the National Institute of Statistical Sciences (under the Federal Grant and Cooperative Agreement Act (1977)). As part of that agreement, three research teams were established. Each team is comprised of a senior faculty member, a more junior faculty member, a post doctoral fellow, a statistics graduate student, and NASS researchers. Notice that this team structure provides the post doc and graduate student an opportunity to become familiar with the agency and for both them and the agency to assess their potential for working there later. It provides a focused research experience for the young NASS researchers, and faculty members have an opportunity to work on meaningful research problems. The problems identified for them to work on are as follows:

Team 1: Multivariate Imputation Mechanisms and Valid Mean Squared Error Estimation: Agricultural Resource Management Survey – Phase III

Team 2: New Design and Estimation Methodologies for Biased Self-Exclusion (Under-coverage): Estimation of Small Farms

Team 3: Statistical Multi-Source Predictive Models and Error Estimates: Major USDA Crop Production Forecasts and Estimates

Although each agency will have its own concerns and issues, these examples illustrate some of the numerous ways to assemble the right set of talents to work on an identified research need.

3. Individual Research

Unless the research problem is extremely well focused, the researcher can benefit from replicating the results obtained from the current methods because, until results can be replicated, the current methods are not fully understood. This is easier if good documentation has been maintained. Replication can take a substantial amount of time. Often current practice is found to deviate, at least to some extent, from the documented procedures. The researcher should understand why the changes were made and update the documentation. In the process of replicating the work, the researcher should note when the reason a particular approach was taken is unclear and consider possible alternate approaches. If necessary, he/she should follow-up with the relevant people to gain a full understanding of what is being done now and why.

Once the potential areas for improvement have been identified, literature reviews help the researcher assess whether a viable solution exists. If not, then identifying the most appropriate approach and extending it to the current problem is often a good strategy. As with any applied statistical research, possible solutions need to be thoroughly tested. Consideration must be given to the ability of the organization to adopt a proposed procedure. For example, if the method requires people to conduct complex analyses that cannot be fully implemented using currently available software packages, then it must be determined whether the agency can realistically implement it.

There are other issues of practicality to consider, as well, such as the potential implications of incorporating a complex statistical algorithm into a production processing system. There is limited processing time available to a government agency in trying to publish survey statistics on a timely basis. Insertion of a procedure that would require excessive processing time to run, even if statistically optimal, could well be impractical. Such considerations sometimes lead to a good approach being found more suitable than an optimal one. Any proposed method must be fully validated and compared to current ones. Then the researcher needs to present the results in both written and oral form. The results should be fully reviewed by colleagues both within and outside the agency, and this might lead to further revisions and improvements. Young researchers often find it difficult to move through these steps alone, and proper mentoring can be critical to their success.

4. Team Research

Research teams are helpful if the problem is especially complex, perhaps requiring diverse sets of knowledge. As mentioned earlier, teams can be used to help build research capacity. From identifying the team leader to ensuring that the right set of talents are present, establishing the team is an important part of the process. Once the team is identified, the team leader needs to work to create an environment where every team member feels comfortable participating fully in all discussions. No one person should dominate and mutual respect should be present. Members should be free to try things that are not standard and that may not work. They should be able to report both failures and successes so that all team members can understand what was learned, so as not to replicate failures, and to build on successes. Although not essential, activities outside of work often lead to an increased sense of team. For example, our NISS-NASS team usually goes to lunch together when we are in one location. We have had dinners together in restaurants as well as the homes of team members. One of our team members had never been to a baseball game, so we went to one of those. Although this did not fall specifically under work and will not be possible in all settings, by getting to know each other better as people, we have come to know each others' personal strengths and weaknesses. We want each other and the team to be successful, and this carries over into work.

The team I led was asked to review the methods associated with estimating the number of farms in the United States based on the June Area Survey, the JAS. This survey is conducted each June. Each member was to serve a role and was assigned an area of responsibility. These can change over time, but having responsibilities helps each member understand that (s)he has an important contribution to make. On our team, the two NASS researchers had the contacts and understanding to help determine the current procedures and to ensure that the team's replication of the work was accurate. One had worked in several areas of NASS before coming to RDD. This gave her a broad understanding of the process, and she knew whom to contact when additional information was needed. She also had a good understanding of the Census of Agriculture, allowing her to match records that no one else on the team had the experience to do. The other was a relatively new recruit to NASS. She has wonderful computational skills and a good research mind. She conducts extensive data exploration to gain a deeper understanding of where data problems may lie. She also replicates each proposed method to ensure that it is translatable into the NASS framework. After about 7 months, the team could replicate the farm estimates obtained by NASS, and a set of improved methods had been proposed. With NASS administrators, the team's research agenda for the upcoming year had been set. Then, in the spring, I got an emergency e-mail from the NASS researchers. They had just become aware that some of the values had been imputed. Because they were fully engaged in all aspects of the work, they recognized this as a potential issue and reported to the full team so that the impact, if any, on the team's project could be assessed.

The graduate student on the team has spent numerous hours gaining an understanding of the literature as it relates to the June Area Survey and has suggested several methodological

improvements. The junior faculty member has a strong background in survey sampling and has led the effort in the development of a proposed follow-up survey for the JAS that builds on prior work conducted within the agency. Finally, a post doctoral fellow was assigned to the team. As sometimes happens, the post doctoral fellow initially chosen for our team was not a good match. By the nature of the problem, the team was working on numerous avenues at the same time. When one person learned something new, it was presented to the group, and others began building upon it. This led to rapid progress in understanding the current methods, but that environment is uncomfortable for someone who is accustomed to working alone from first principles. She did not feel successful, and others on the team were concerned that she would not be able to do the work that was needed. She decided to find a different job, one for which she is well suited. A new post doctoral fellow was sought. In addition to the NISS and NASS administrators, all team members participated in the selection of the new post doctoral fellow, and the final choice was a unanimous decision. That said, the team had been together for more than 6 months when the new post doc joined, and it took time for him to learn what the team had done and to find his place. The transition was obviously complete when he told the graduate student that he thought an important source of variation in a variance computation had been omitted. This led to a discussion that improved the work. Now he is focusing on estimates and standard errors at the state level.

The team moves through the research process, much as an individual would. However, as indicated earlier, by building on the individual strengths of the team members, a team should be able to make more rapid progress toward the solution. Also, I have found it useful to have an open dialogue with NASS administrators so that the team is always clear as to agency priorities and how its work will address at least some of these. Again, communicating the results both orally and in written form is important.

4. Discussion

In some government agencies, researchers write reports that are put on a shelf so that they will be available if the need arises. Although this type of research has its place, it does not provide the strong sense of responsibility and motivation that the research within NASS affords. The research questions are designed to address agency needs. Those conducting research know that, if they provide new methodology that addresses the questions and can realistically be implemented, their work will be incorporated into NASS procedures. This is an exciting and wonderful opportunity. Many researchers work their whole career without ever having such an impact. The idea that the research will actually make a difference in how things are done in the future has certainly been one of the strongest motivational forces for our team and for researchers within NASS. This would not be possible without strong support from upper administration.

REFERENCES

The Council on Food, Agricultural and Resource Economics (2007) *Improving Information about America's Farms and Ranches: A Review of the Census of Agriculture*. Washington DC.

Available on the web:

http://www.cfare.org/publications/20070307cfare_census_review_Full_Report.pdf.

The Council on Food, Agricultural and Resource Economics (2009) *A Review of the USDA_NASS Agricultural Prices Program*. Washington DC. Available on the web:

http://www.cfare.org/reviews/CFARE_NASS_FULL_BOOK_email.pdf

Federal Grant and Cooperative Agreement Act (1977) Available on the web:

<http://thomas.loc.gov/cgi-bin/bdquery/z?d095:HR07691:@@L&summ2=m&>

National Research Council (2007) *Understanding American Agriculture: Challenges for the Agricultural Resource Management Survey*. The National Academies Press: Washington DC.

Available on the web: http://www.nap.edu/catalog.php?record_id=11990