United States Department of Agriculture

Economic Research Service

Economic Research Report Number 236

August 2017

Economic Experiments for Policy Analysis and Program Design: A Guide for Agricultural Decisionmakers

Nathaniel Higgins, Daniel Hellerstein, Steven Wallander, and Lori Lynch

United States Department of Agriculture

Economic Research Service
www.ers.usda.gov

Download the charts contained in this report:

- Go to the report's index page
- Click on the bulleted item "Download ERR236.zip"
- Open the chart you want, then save it to your computer

Recommended citation format for this publication:

Higgins, Nathaniel, Daniel Hellerstein, Steven Wallander, and Lori Lynch. *Economic Experiments for Policy Analysis and Program Design: A Guide for Agricultural Decisionmakers*, ERR-236, U.S. Department of Agriculture, Economic Research Service, August 2017.

Cover image: iStock

To ensure the quality of its research reports and satisfy government-wide standards, ERS requires that all research reports with substantively new material be reviewed by qualified technical research peers. This technical peer review process, coordinated by ERS' Peer Review Coordinating Council, allows experts who possess the technical background, perspective, and expertise to provide an objective and meaningful assessment of the output's substantive content and clarity of communication during the publication's review.

United States Department of Agriculture

Economic Research Service

Economic Research Report Number 236

August 2017

Economic Experiments for Policy Analysis and Program Design: A Guide for Agricultural Decisionmakers

Nathaniel Higgins, Daniel Hellerstein, Steven Wallander, and Lori Lynch

Abstract

Economic experiments can help inform the design and implementation of government policies, especially newly conceived or novel policies. This report reviews experimental methods, with a focus on when and how they can be effectively used to inform agricultural program decisions. To illustrate the capabilities of experimental methods, five case studies are presented. First, a laboratory study examines whether it is possible to improve the cost-effectiveness of auctions similar to those used in voluntary land retirement programs. A second laboratory study, with both students and farmers, illustrates how a combination of targeting and bonuses may be able to improve environmental outcomes by encouraging coordination among conservation program enrollees. The third study, conducted in the field to more closely mimic real-world conditions, measures how farmers trade off current income against future income and finds farmers will accept less money if paid today rather than in the future. The fourth and fifth studies are randomized controlled trials testing the marketing of USDA programs; these show how outreach letters can increase participation in the Conservation Reserve Program and in county committee elections.

Keywords: behavioral economics, Conservation Reserve Program, conservation targeting, randomized controlled trials, economic experiments, laboratory experiments, field experiments

Acknowledgments

The authors would like to thank Constance Newman and Jacob Fooks, U.S. Department of Agriculture (USDA), Economic Research Service (ERS); and Richard Iovanna, USDA, Farm Service Agency, for data and advice. They would also like to thank the following individuals for technical peer reviews: David Just, Cornell University; Brian Roe, Ohio State University; John Pender, USDA/ERS; Joy Harwood and Richard Iovanna, USDA, Farm Service Agency; and Noel Gollehon, James Rowe, and Janet Perry, USDA, Natural Resources Conservation Service. We also thank Dale Simms and Cynthia A. Ray, USDA/ERS, for editorial and design services. Nathaniel Higgins was formerly with ERS when this research was conducted.

Contents

United States Department of Agriculture

A report summary from the Economic Research Service

August 2017

Economic Experiments for Policy Analysis and Program Design: A Guide for Agricultural Decisionmakers

Nathaniel Higgins, Daniel Hellerstein, Steven Wallander, and Lori Lynch

What Is the Issue?

For many policy design questions—such as setting up enrollment structures, establishing program eligibility criteria, investing in education and outreach materials, and ranking/selecting applications—policymakers often seek evidence on how well alternative approaches will achieve program goals. Experiments offer a way to gather data for these uses. This report examines the use of economic experiments for building evidence to inform policy, providing an introduction to the field of experimental economics and discussing the increasing use of field experiments and randomized controlled trials (RCTs) in the social sciences and in government. We present several case studies illustrating recent uses of experimental methods to inform agricultural policies.

What Did the Study Find?

This report provides a number of lessons gleaned from both the growing literature on experiments and from actual experiments to examine ways of improving agricultural programs.

- ***Laboratory experiments***—often with students using networked computers—are useful for testing basic theory and initial evaluation of prospective policies where naturally occurring data do not exist. Lab experiments offer a degree of control, allowing researchers to change one factor at a time to determine causality.
- In contrast, ***field experiments***—using people in a real-life environment—offer less control but have a "context" that more likely mirrors the program's target environment with participants who better represent the population of interest.
- The most credible tests of the likely impact of a policy or policy instrument occur when two similar groups are compared—setting one as a "control" and the other as the "treatment" group. In particular, experiments based on ***randomized controlled trials (RCTs)***—often large-scale field experiments where subjects are randomly assigned to control and treatment groups—can provide evidence that is gathered in a manner that isolates causality. There is a long history of RCTs being used to inform Federal policymaking, particularly in education, welfare, and health care.
- Lab experiments are relatively inexpensive and easy to perform. Field experiments—especially those evaluating key aspects of large programs—are more expensive and difficult to administer. A hybrid, cost-effective approach is to test new policy mechanisms iteratively, moving those that do well in laboratory experiments into a field setting, such as via a "pilot" program.
- Experimentation need not be limited to wholesale tests of new policies and programs. ***Administrative experiments***, which systematically test new ways of operating an ongoing program, can provide low-cost and rapid information to improve program operations. A program's administrative data can be used by the experimenter to readily measure the effect of a treatment without the burden and expense of additional surveys or focus groups. For example, in the "Encouraging Participation in the Conservation Reserve Program"

experiment discussed below, researchers used administrative data to identify which new enrollees were part of the experimental treatment.

- Opportunities for experimentation, including the use of administrative experiments, occur frequently within the normal policy process and include: (1) when a new program is rolled out or during a new iteration of a program (such as a yearly program signup); (2) when a program is oversubscribed and decisionmakers must choose which applicants to enroll; (3) when a program's target populations or goals change; and (4) when a program is undersubscribed and interventions (such as new methods of outreach) may increase participation.

To illustrate the use of a variety of experimental approaches and to highlight key experimental design features, this report reviews the methods and findings of five case studies where experiments were used to investigate issues pertinent to agricultural policy design.

- "**Asymmetric Auctions and the Design of Enrollment Mechanisms**." This *laboratory experiment* investigated whether it is possible to design a more cost-effective means of enrolling participants in voluntary conservation programs. Researchers presented students with a competitive enrollment mechanism—a *reverse auction*—wherein a purchaser (such as a conservation program manager) accepts a number of offers from a variety of potential program participants; the purchaser has limited information about these potential participants. The costs of operating the hypothetical program were reduced by as much as 14 percent when information about each potential participant is built into the enrollment mechanism.
- "**Do Bonuses Increase the Agglomeration of Buffers Along Streams**?" This *laboratory experiment* investigated how to coordinate enrollment in a voluntary program when overall benefits depend on the geographic distribution of participants. Using both student and farmer participants, researchers investigated whether paying participants "spatial bonuses" (for enrolling near neighbors), using "spatial targeting" (prioritizing adjacent fields for enrollment), or both can improve environmental outcomes. Spatial targeting was found to increase total environmental benefits and social welfare, but spatial bonuses did not.
- "**Personal Discount Rates**." This *field experiment*, conducted with farmers, investigated the tradeoff between upfront and future payments. Participation in conservation programs may be affected by this timing; farmers who strongly prefer upfront payments might pass on a program offering an income stream spread over time. Researchers presented farmers with an option of receiving a payment of several hundred dollars now or a higher payment 9 months in the future. Farmers in this study, on average, needed to be paid an interest rate of more than 28 percent before forgoing the upfront payment for the delayed payment.
- "**Encouraging Participation in the CRP**" and "**Encouraging Voter Participation in FSA County Committee Elections**." These two *administrative experiments* were structured as randomized controlled trials to study the effectiveness of farmer outreach initiatives used by the USDA's Farm Service Agency. Researchers found that simple outreach—customized program re-enrollment letters in the first, and customized voter reminder postcards and ballots in the second—increased farmland owner interest in participating in a conservation program and increased voter participation in county committee elections. The reminder letters generated one additional offer to re-enroll in a conservation program for a printing/mailing cost of $41; likewise, the printed outreach resulted in an additional vote for every $28.55 spent.

How Was the Study Conducted?

This report consists of three parts. First, we review fundamental research in the methodology of experimental economics and experiments in social science more generally. We use this review to establish a small set of "common practices" that researchers use when conducting economic experiments. We next examine successful instances when public policy was informed by experimental studies and analyze the characteristics of these influential studies. Finally, we synthesize several economic experiments conducted by ERS and collaborators. Some were conducted in the laboratory, some in the field, and some in partnership with USDA program agencies. These case studies illustrate how experiments can address real policy questions and provide useful insights.

Economic Experiments for Policy Analysis and Program Design: A Guide for Agricultural Decisionmakers

Introduction

> "It is a commonplace that, in its choice of method, economics is limited by the fact that resort cannot be had to the laboratory techniques of the natural sciences."
>
> – Chamberlin (1948). *An Experimental Imperfect Market*

In fields of scientific inquiry such as physics, biology, and crop science, experiments are routinely conducted to test theories and advance knowledge. Galileo is said to have conducted an experiment in which he dropped two balls—each made of the same material but having different masses—off of the Leaning Tower of Pisa to prove that all objects fall at the same rate.[1] James Lind conducted an experiment in 1747 where sailors with scurvy were given otherwise identical rations augmented by a different "acidic" food or beverage. Those who received citrus recovered; eventually it was discovered that scurvy was caused by a lack of Vitamin C.[2] In the early 20th century, Ronald Fisher developed new methods of experimentation in an effort to determine what caused crop yields to vary so substantially from year to year.[3] The experiments he designed helped to establish the importance of fertilizer and weed control in enhancing yields.

In contrast to this long history of experimentation in other fields, until recently the prevailing view was that economics is not an experimental science. Without economists engaging in experimental research, most policy research relating to economic behavior has historically been done without the benefits—or costs—of experimental methods.

Creating experimental tests of important behavioral and economic phenomena can be challenging. For example, it is difficult to test the effect of interest rates on inflation, or the effect of crop insurance premiums on national corn production. In each case, the outcome under investigation—inflation or corn production—is the result of complicated and ever-changing interactions among millions of people. At the very moment we observe the level of inflation or corn yield, the conditions that created them are changing.

Recent developments have, however, raised confidence in experiments as a useful tool for economic and policy analysis. Over the last 30 years, economists have increasingly identified areas of economic behavior and policy design that are amenable to experimental analysis. In the early 1980s, the number of articles in economics journals using an experimental method averaged around 50 per

---

[1]The idea of the experiment is attributed to Galileo, but whether Galileo actually carried out the experiment or instead simply conceptualized the experiment—a so-called *thought experiment*—is a matter of historical debate. See, for example, Palmieri (2009).

[2]See Hughes (1975), for example.

[3]Fisher wrote extensively about both statistical methods and crop science. See *Statistical Methods for Research Workers* (1925) and *The Design of Field Experiments* (1935) for examples of the former, and S*tudies of Crop Variation I, II, IV, VI*, (1921, 1923, 1927, and 1929) for examples of the latter.

year. By the late 1990s, this rate had quadrupled (Holt, 2006; Levitt and List, 2007). Evidence of the maturity of experimental economics as a discipline include Vernon Smith's Nobel Prize in 2002 for his work in the field,[4] the establishment of influential journals dedicated to experimental work,[5] and the extensive use of field experimentation by leading researchers in many fields.[6]

Beyond their role in reshaping how economic research is conducted, social science experiments have been used extensively to influence the design of both markets (especially new markets that operate on the internet) and public policies. More than a dozen Federal agencies have used experiments to develop new programs or better understand policy. Experimental economic methods have played important roles in the successful design of spectrum auctions carried out by the Federal Communications Commission (Klemperer, 2002; Guala, 2001), in convincing the Federal Aviation Administration to propose auctioning rights to land and take off from crowded airports (Ball et al., 2007), and in the Securities and Exchange Commission determination of whether "uptick" or "short-selling" restrictions should be eliminated (Office of Economic Analysis, 2007).

As the practice of experimenting to answer economic and behavioral questions has evolved in recent years, some best practices have emerged, and experiments have taught economists and other social scientists important lessons.

[4]Vernon Smith was awarded the prize "for having established laboratory experiments as a tool in empirical economic analysis, especially in the study of alternative market mechanisms." Smith shared the prize with Daniel Kahneman, who was honored "for having integrated insights from psychological research into economic science, especially concerning human judgment and decisionmaking under uncertainty."

[5]For example, *Experimental Economics* began publishing in 1998 and now has an impact factor—a measure of professional influence—higher than the *Journal of Econometrics*.

[6]For example, see List and Rasul (2010) and Banerjee and Duflo (2008).

What Is an Experiment?

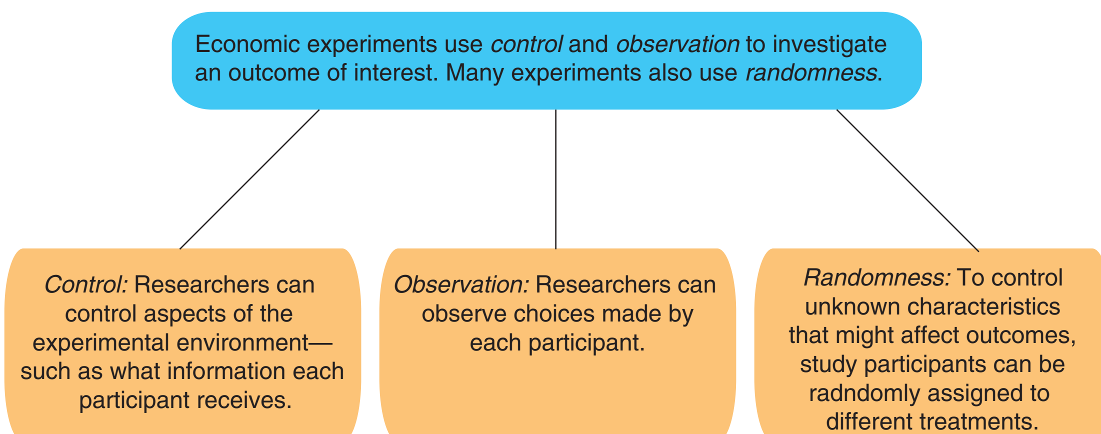
The defining characteristics of experimentation in any scientific field are the use of *control* and *observation* to investigate an outcome of interest. Additionally, certain kinds of experiments use *randomness* as a fundamental part of their design (figure 1).

A clinical trial, for example, works this way: a group of individuals is assembled; some of those individuals are randomly selected to receive a new medicine, while others receive nothing or a placebo (a sugar pill); finally, the health of all individuals is measured and compared. Health differences are presumed to be due to the new medicine. In this case, researchers can ask participants to avoid strenuous exercise during the study (*control*) if doing so makes it more difficult to determine the effect of the new medicine on health. It is critical that researchers are able to *observe* a sufficient number of individual outcomes to conclude with confidence whether or not the new medicine impacted average health.[7] *Randomness* dictates that the assignment of each study participant to receive either the new medicine or a sugar pill occurs randomly, as if by a coin flip.

By using control and multiple observations, experiments allow researchers to interpret whether the treatment applied *causes* the outcome.

Experimental methods in economics amount to observing choices made by people. When experiments are carefully structured, these choices can indicate what drives decisions. These drivers can include prices and attributes of goods and services, as well as behavioral factors such as susceptibility to how information is presented, social norms, and cognitive burden.

Figure 1
Control, observation, and randomness in economic experiments



Source: USDA, Economic Research Service.

[7]The simplest experiment would observe health outcomes of just two individuals. However, because many factors influence health and because researchers cannot measure and control all of these factors, an experiment with such a small sample size is unlikely to reveal the effects of a new medicine.

Experimental methods often complement other tools. For example, simulation models can be used to predict how alternative policy interventions might lead to a variety of impacts that cascade through markets. *Econometric studies* can capture how personal and environmental characteristics can affect a variable of interest, given data with sufficient variation and a properly specified model. *Contingent valuation and choice experiments* can provide information on the social welfare impacts of alternative policies. All of these economic methodologies require some degree of assumption about how economic phenomena occur and how humans behave. Economic experiments cannot replace the breadth and flexibility of these methods, but in many cases, experiments give researchers and policymakers the ability to test these assumptions. In some cases, experiments reveal outcomes not predicted by these other methods.

Experiments in Economics

There are many types of experiments conducted by economists and other social scientists, each with distinct features and roles to play in informing policy. Roth (1986) suggested organizing experimental studies by purpose, paraphrased here:

(1) experiments designed to test theory,

(2) experiments used to establish new facts that theory must account for, and

(3) experiments designed to inform policy.

There is a long tradition of experiments to test basic academic theories in both psychology and economics (see box, "Early Economic Experiments"). These experiments most often take place in controlled settings ("the lab"), with feedback determining which theories should be refined and which should be discarded. For example, experiments have compared how well two different theories explain the way that individuals behave in financially risky situations.[8]

Experiments can spur the creation of new theories as well as test existing ones. For example, experimentalists found that people participating in auctions in a classroom laboratory frequently bid much higher than predicted. This recurring observation has led theorists to develop and test many possible explanations, allowing for better predictions of real world behavior.[9]

[8]For example, Camerer (1989) conducted an experiment using choices between lottery tickets to compare *expected utility theory and prospect theory*. See Barberis (2013) for a recent review of other work on this subject.

[9]Explanations for overbidding include risk aversion (Cox et al., 1988), weak incentives (Harrison, 1989), and behavioral theories such as the effect of regret on bidding (Filiz-Ozbay and Ozbay, 2007).

Box 1

Early Economic Experiments

Perhaps the first experiment conducted by an economist was by William Jevons in 1870. Jevons' study actually consisted of three experiments. The first investigated the relationship between a weight thrown and the amount of work done, where work was measured by the total weight thrown times the distance it was thrown. The other experiments were similar. Each considered the best way to do physical work—should work be done in fewer, more rigorous movements or in more, less rigorous movements—given the constraints of the human body. In this unique study, Jevons was both the experimenter and the participant: he threw all the weights himself. If Jevons had simply thrown each weight a single time, he might have been concerned that the results were due to chance—that a gust of wind aided some of this throws, for instance. By repeating each exercise many times, Jevons was able to ensure that his results were due to the tradeoffs inherent in physical work.

While Jevons' study used experimental methods, it would not be recognizable as a modern economics experiment—it did not include any measure of behavior and was generally thin on what we recognize as economics in 2017. In what most people identify as the first modern economic experiment, Chamberlin (1948) constructed a market economy in a classroom (students bought and sold playing cards that Chamberlin gave to them), allowing the behavior of buyers and sellers to be monitored. Chamberlin could observe the process of price adjustment and bargaining that leads to market equilibrium, a process that could not be observed outside of a laboratory setting.

The Lab and the Field

Experiments designed for the first two of Roth's three purposes are often conducted in laboratory settings, typically classrooms with networked computer terminals and workstations for each participant. A hallmark of the laboratory setting is the ability to vary a single factor, holding all other factors fixed. With this *control,* any observed differences across treatments can be convincingly attributed to the single, varying factor.

The ability to control the laboratory environment makes lab experiments particularly well suited to testing and refining theory or the fine details of policy design. However, this control comes at a price.

A laboratory environment is, by necessity, artificial.[10] Observing behavior in a lab is convenient and affords a high degree of control over the environment, but the context[11] of the laboratory and the participants are in many ways dissimilar to those in naturally occurring environments. Economic and government policy—including diverse questions such as how to most efficiently enroll farmers in a conservation program, or what foods to include in a food assistance program—is formulated in an environment that is much less controlled, and more dependent on context and ever changing variables, than a lab environment. Because of this, economists, political scientists, and other social scientists have developed techniques to conduct experiments in natural settings, or "the field" (see box, "Field Experiments," for the hallmarks of a field experiment).

Ultimately, both lab experiments and field experiments attempt to identify the effect of a treatment—whether a simple change in how people communicate, a tweak to a government program, or a whole new policy—on an outcome of interest (figure 2).

Box 2

Field experiments

The term *field experiment* refers to any experiment that is not conducted in a classroom laboratory using student participants, so variation between experiments is wide. Harrison and List (2004) introduced qualifying terms that have since become standard in economics.

An ***artefactual*** field experiment takes place in a laboratory but includes participants other than students. Conducting a lab-style experiment with farmer participants rather than student participants would constitute an artefactual field experiment.

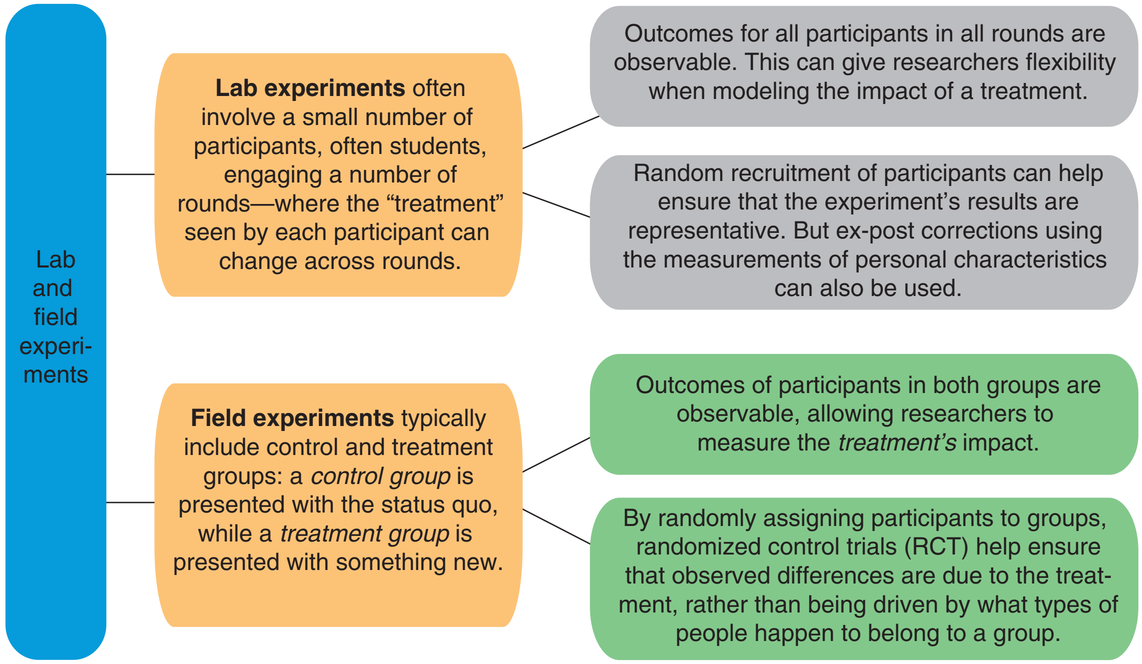
A ***framed*** field experiment is an experiment in which the context of the experiment is specifically modeled after the natural context in which the phenomenon under study takes place.

A ***natural*** field experiment is one in which the experimenter does not disturb the natural environment in any way. Randomly assigning farmers to receive one of two different letters, then observing what impact the different letters have on the likelihood that a farmer signs up for a government program, constitutes a natural field experiment.

[10]See Bardsley et al. (2010) for a discussion of artificiality in economic experiments. An experiment that takes place in an artificial environment is said to have low "ecological validity."

[11]As defined in the report's glossary, "context" refers to the situation that an experiemtn is conducted in.

Figure 2
Lab and field experiments



Source: USDA, Economic Research Service.

Lab and field experiments accomplish this task differently, however. While laboratory experiments impose a high degree of control in an attempt to eliminate factors that might confound causal interpretation, field experiments emphasize the use of randomness to balance these factors.

Crucially, in a properly designed field experiment, the only systematic difference between the treatment group and the control group is the application of the treatment. A pool of experimental subjects typically has many characteristics—such as age, health, or ethnicity—that may influence the outcome of interest. Randomized control trials (RCTs) are often used to account for this heterogeneity. By randomly assigning subjects to control and treatment groups, RCTs balance these potentially confounding factors between the groups, allowing effects due solely to the treatment to be readily detected (see box, "Randomization in Lab and Field Experiments"). As an example, random assignment in a clinical trial is meant to ensure that the group receiving a new medicine is unlikely to be younger and healthier from the start than the group receiving a placebo.

Field experiments are conducted in the participants' natural environment and do not disturb that environment, conferring some practical advantages. For one, phenomena can be observed that would be difficult or impossible to replicate in the lab. For example, there are many reasons why individuals or businesses might not behave in the same way in a lab as in their daily activities, one of which is the desire not to be detected engaging in certain behaviors.[12] Also, the incubation between field study findings and full-scale policy implementation is likely shorter than that between lab findings and policy. Both field experiments and lab experiments are useful and can be used in concert to inform policy.

---

[12]For example, field experiments have helped to identify discrimination against minorities in economic environments (List, 2004), including housing markets (HUD, 2012).

Box 3

Randomization in Lab and Field Experiments

The three signature features of experiments are randomness, control, and observation. Both lab and field experiments impose some degree of explicit control (lab experiments more, field experiments less) and rely on sufficient observation of behavior to draw statistically relevant conclusions. Lab and field experiments, however, typically use randomness differently. Lab experiments might not include any random assignment of individuals to treatments. In fact, a well-designed lab experiment might consist of a single sequence of treatments or "games" that each participant faces—so that each participant may face the same set of treatments (albeit at different times). In this case, the role of randomness is simply to select a representative set of individuals from a population. That is, to draw general conclusions from a lab experiment, one must assume that the individuals participating in the experiment represent a random sample from a particular population, or one must use a modeling strategy to match the observed characteristics of the sample to the characteristics of the overall population.

For field experiments, randomization is used to control for confounding factors—in particular, via the use of randomized controlled trials (RCTs). The term RCT is most often associated with the field of medicine but is increasingly used by social scientists to refer to certain field experiments. While the terms *field experiment* and *RCT* are sometimes used interchangeably to mean "an experiment taking place outside the lab," RCTs are more often associated with large-scale tests within existing institutions. The distinction between the two terms is not precise, and many practitioners do refer to a single experiment using both terms. See Glennerster and Takavarasha (2013) for an introduction to RCTs in the social sciences. In this report, RCTs and field experiments can be thought of interchangeably.

What Is a "Policy-Relevant" Experiment?

While experiments can be relevant to policymakers to varying degrees, *policy experiments* are meant to deliver evidence that is actionable. For example, the Oregon Medicaid experiment and the Health and Human Services Head Start experiment both use randomized controlled trials to evaluate large social programs (see box, "Two Examples of Large Experiments"), enabling direct assessment of program activities and benefits. However, experiments can also be used to test ways to improve programs without conducting a programwide evaluation. *Administrative experiments*—pilot tests embedded directly into administrative operations—test facets of programs rather than whole programs themselves.[13]

Of course, most experiments exist along a continuum, from pure tests of abstract theories to pure policy experiments. Like field experiments, lab experiments can be policy oriented too. And although testing basic theory is the most common use of classroom-style lab experiments, they can also serve as a *demonstration* tool when conducting tests of new mechanisms (a practice sometimes referred to as *test-bedding*).

Lab experiments can be helpful in demonstrating concepts to nonspecialists in several ways. For example, by participating in experiments themselves—a prospect made possible by the simple setup of many lab experiments and the widespread use of mobile technology—program managers and other individuals who are not expert researchers can better understand the findings from research. For example, during his term on the President's Council of Economic Advisers, Professor John List conducted an experiment using senior officials in the George W. Bush administration as participants. List used the experiment to illuminate the likely outcome of an auction proposed under the Clear Skies Act of 2003.[14]

Second, the quantitative results generated by lab experiments can often demonstrate to nonspecialists why one particular policy or program element is preferred in a way that is straightforward, easy to interpret, and less reliant on specialized background than theoretical arguments or arguments based on naturally occurring phenomena. In both this case and instances of demonstration-by-participation, the experiment serves as a communication tool rather than a research tool.

Test-bed experiments are designed to create "working prototypes of a process that is going to be [or could be] employed in a complex environment" (Plott, 1997). The laboratory can be a particularly good place to test out new mechanisms, whether whole new institutional designs or changes to institutional rules. When a new market or program mechanism is proposed, there is no way to test this mechanism using an administrative experiment because the proposed design has not yet been implemented. Using the laboratory provides a preliminary assessment of a proposed mechanism before it is rolled out or scaled up and offered to individuals and businesses.

When the Federal Communications Commission decided to auction licenses for personal communications systems for the first time, rather than granting licenses using an administrative procedure (Plott, 1997; Cramton 1995; Ausubel et al., 1997), lab experiments were used to test the proposed design of the auction. Such details could influence the eventual ownership of the valuable licenses as

---

[13]The term "administrative experiment" was, to the best of our knowledge, first used by Simon and Divine (1941). More recently it has been used in a fashion similar to the way we use it here by Cody et al. (2015).

[14]Used with the permission of John List (personal communication, 3/07/2017).

well as the price paid to the Government for those licenses. Conducted in simple laboratory settings and often using students as the participants, these experiments were important first empirical tests of the proposed auction rules. While conducting test-bed experiments does not guarantee program success, it allows for trial and error and is a useful check on unintended consequences before being applied to high-stakes programs.

Box 4

Two Examples of Large Experiments

In 2008, Oregon expanded coverage of its Medicaid program to an additional 30,000 applicants. Since 90,000 people applied for 30,000 application slots, the program needed to ration the service. The selection of participants took place by random lottery, in effect creating a randomized controlled trial (RCT) experiment where the treatment was the ability to participate in Medicaid. Crucially, the lottery was both a neutral way to ration a service and a method that supported research. The only thing that differed systematically between the treatment group (those individuals who won the lottery and were randomly selected to receive coverage) and the control group (those individuals who did not win the lottery) was winning the lottery. We can surmise that any difference in health outcomes across the two groups was due to access to Medicaid. In this case, the experimental environment and the target environment were the same.

Baicker and colleagues (2013) found that Medicaid coverage generated no significant differences in measured physical health outcomes in the first 2 years between those who won the lottery and those who did not. However, access to Medicaid coverage did increase use of healthcare services, raise rates of diabetes detection and management, lower rates of depression, and reduce financial strain, all of which may improve longrun health. The Oregon policy field experiment provided policymakers with confidence that it was the expansion of Medicaid (and not other external factors) that was responsible for higher rates of diabetes detection and management, for instance.

While this experiment is capable of teaching us important lessons about the behavior of those who applied for Medicaid, it might have less to say about the benefits that might accrue to individuals who might qualify but did not choose to apply—there may be selection *effects*. Furthermore, extending these findings to other contexts (say, Medicaid in Alabama) is not automatic—the study's external validity is unknown.

A second example of an RCT experiment of a large Government program is the congressionally mandated Head Start Impact Study in 1998. This study aimed to determine the impact of participation in Head Start on "children's school readiness and parental practices" (U.S. Department of Health and Human Services, 2005). HHS conducted the study, which randomized almost 5,000 applicants to oversubscribed Head Start programs, determining entry into the programs by lottery. HHS released a final report in 2010, finding multiple positive impacts on school readiness of 3- and 4-year-olds at the end of their Head Start year. And at the end of 1st grade, Head Start students scored higher on vocabulary tests. However, for both 3-year-olds and 4-year-olds, math skills, pre-writing, children's promotion, or teacher assessments of children's school accomplishments or abilities were not significantly impacted by Head Start after kindergarten or 1st grade.

The Importance of Validity: Can Experimental Economics Inform the "Real World?"

Many have expressed concerns about the artificiality of laboratory experiments (see box, "Hypothesis: Can Experiments Elicit Meaningful Responses?" and Guala, 1999). These same concerns lead to questions about the policy lessons that can be drawn from experiments. A program manager must determine what can be learned from an experiment (or any research) that takes place in an environment that differs in some way from his or her policy environment. In short, does the experiment have *external validity?*

The validity of an experiment often depends on the extent to which the results are considered convincing; scientists often evaluate the validity of a given experiment on a sliding scale that is a matter of professional opinion rather than rigid definitions. An experiment meant to test how a treatment affects the behavior of commodity prices might be more persuasive if enacted on a mock

Box 5

Hypothesis: Can Experiments Elicit Meaningful Responses?

Can economic experiments reveal information on actual human behavior? Beginning in the early mid-20th century, a debate on this question began and it continues to shape current practice.

In 1931, L.L. Thurstone (1931) offered each of his participants a series of choices over bundles of goods—would they prefer eight hats and eight pairs of shoes, or six hats and nine pairs of shoes, for example. Thurstone used the choices to map out the relative value of each of the goods and captured the tradeoffs each person was willing to make to obtain more hats, shoes, or coats. Thurstone's experiment concluded very little. The methods that Thurstone used, however, spurred a debate that had important consequences for the current practice of experimental economics.

In 1942, Wallis and Friedman criticized Thurstone's study, contending that Thurstone's method was incapable of revealing his participants' true preferences. The primary difficulty with the method was that Thurstone presented his participants with artificial and hypothetical choices. Thus, when Thurstone asked an individual whether he or she preferred to receive eight hats and eight pairs of shoes or six hats and nine pairs of shoes, the participant's answer could not be trusted. Wallis and Friedman did not believe that participants lied when they said they preferred six hats and nine pairs of shoes, but rather Wallis and Friedman believed that participants could not know how they would react to such a choice because the choice itself was encountered inside the artificial world of the experimenter.

Basically, Wallis and Friedman questioned the external validity of Thurstone's experiment. They doubted Thurstone's—or anyone else's—ability to create the necessary conditions for participants to know what they would do if faced with this choice in a *real* situation.

This criticism of an artificial context remains a common complaint lodged against experimental economics.

trading floor rather than in an ordinary classroom, and even more persuasive if it took place on the floor of the Chicago Board of Trade.[15]

Often, a given line of research will proceed from early studies with lower validity that explore new aspects of behavior and program design to a new set of studies that may be narrower in focus but use methods with greater externally validity. There are dozens of other ways to evaluate validity, with specialized terminology to match.

Two notions of validity are of special interest to scientists and policymakers: internal validity and external validity. Program managers and scientists can evaluate both types of validity when presented with experimental results, or when thinking of implementing experiments in their own programs. *Internal* validity is concerned with the identification of causal effect—if A happens, then B happens—in the specific instance of the experiment. An experiment with a high degree of control is more likely to produce results that are internally valid (see box, "Internal Validity: An Example of Competing Explanations"). *External* validity is concerned with the ability to generalize from the experiment to target environments.[16] The box, "Experiment Characteristics," summarizes how internal/external validity as well as control and context vary over several common classes of experiments.

Internal validity and external validity are sometimes in conflict (Roe and Just, 2009). Increasing *control* (narrowing focus) can improve internal validity but may limit the generalizability of findings. The design of an effective policy experiment balances control with generalizability. In practical

Box 6

Internal Validity: An Example of Competing Explanations

Consider a simple experiment to test whether people demand less of a good as the price of the good increases. One way to test this basic assumption is to offer a commodity for sale to many individuals at varying prices, say $10 and $20. Charging different prices to different participants constitutes a treatment. If less of the commodity is purchased at a price of $20 than at a price of $10, the result provides evidence that demand decreases when price increases.

Of course, other reasons may explain a change in demand. Suppose that the commodity used in the experiment was ice cream. The $20 ice cream was offered on a 50-degree day, while the $10 ice cream was offered on a 90-degree day. In this case, the observed change in demand could have been caused either by the treatment (the price change) or by the temperature change. Instead of explaining the impact of price on quantity demanded, the experiment may be demonstrating that people have different preferences for ice cream on 50-degree days and 90-degree days. An experiment that offered $20 ice cream on a 50-degree day and $10 ice cream on a 90-degree day would be a poorly designed experiment with low internal validity. The experiment would have higher internal validity—and would be more convincing—if the temperature were the same during both the $10 and $20 treatments.

When an experiment is well-designed and has high internal validity, the presumption is that a change in the outcome—quantity demanded—is caused by changes in the treatment: price.

[15]When an experiment takes place in an environment that approximates real world conditions it is said to have high ecological validity.

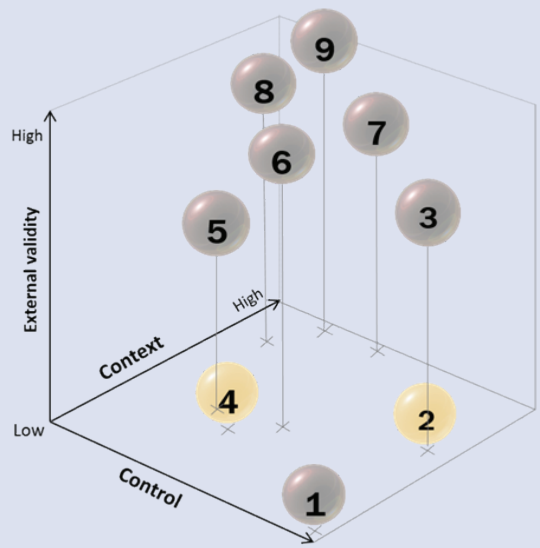
[16]Both terms seem to have been coined by D.T. Campbell and his coauthors (1957, 1969) in a series of papers. Campbell was a psychologist who wrote extensively on methodological issues.

Box 7

Experiment Characteristics

Experiments can be classified along a number of dimensions, including "control," "context," and "external validity." The graphic below considers several common classes of experiments, displaying where they lie along these dimensions; the table lists examples (some discussed in this report) of these classes of experiments.

Box figure 7.1



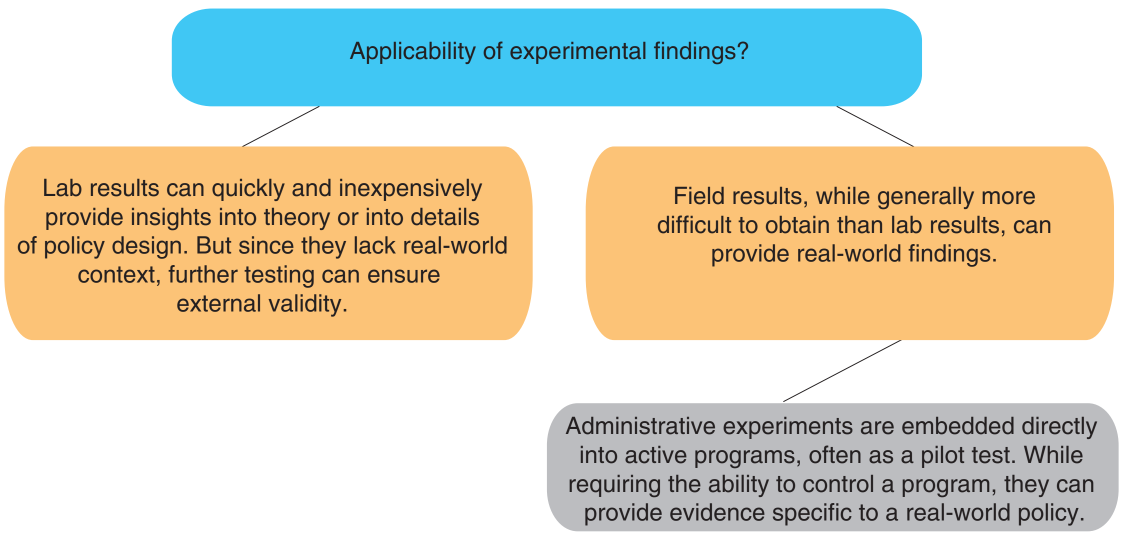
Note that bubbles 2 and 4 float just above the context/control horizontal plane, and are colored transparently so as to reveal the markers of the locations on this plane of bubbles 3 and 5.

Adapted from Messer, Duke, and Lynch (2014)		
	Type of experiment	Example
1	Lab experiment with students, using artificial example	Case study 3a (cost effectiveness of quota auctions, using students)
2	Lab experiment with students, using realistic example	Case study 3b (addressing contiguity in land retirement auctions, using students)
3	Lab experiment with market participants (such as farmers), using realistic example	Case study 3b (addressing contiguity in land retirement auctions, using farmers), and consumer goods under threat of destruction (Messer and Borchers, 2015)
4	Lab experiment with students, using actual goods	Charitable giving for conservation (Zarghamee et al., 2016)
5	Lab experiment with market participants, using actual goods	Responses to food labeling (Council for Agricultural Science and Technology, 2015)
6	Field experiment with market participants	Case study 3c (discount rates of farmers)
7	Field experiment with market participants, using agricultural example	Effects of information on water quality protection (Ellis et al., 2016)
8	Natural experiment, using sample and administrative data	Case study 3d (increasing Conservation Reserve Program offers using nudges)
9	Natural experiment, using full population and administrative data	Case study 3e (increasing participation in county committee elections)

terms, cost and other concerns may favor experimental designs with high control and high internal validity, but results may not reflect or inform important features of the policy of interest (figure 3). For example, the use of students in labs can facilitate experiments (see box, "Using Students as Experimental Participants: When Does It Matter?"). Though researchers have found that students and nonstudents perform similarly in a variety of situations, using students may decrease the external validity of the experiments. As such, researchers may wish to test how well the students' results translate to other populations.

A well-rounded approach to using policy experiments can include initial lab tests using students or other participants, followed by field implementations of the policy to test its effectiveness in context.

Figure 3
Applicability of experimental findings



Source: USDA, Economic Research Service.

Box 8

Using Students as Experimental Participants: When Does It Matter?

A number of researchers criticize the frequent use of students as participants in lab experiments. What do students have to teach, the argument goes, about the behavior of nonstudents? Aren't students different? Researchers have studied the difference between student and nonstudent behavior in a variety of experimental settings (Belot et al., 2015). If the purpose of an experiment is to test basic economic theory, existing studies suggest it does not matter whether the experimental participants are students (Crosson, 2005). Moreover, students tend to perform similarly to nonstudents in a variety of situations, including when students might be expected to behave very differently (Frechette 2015; Exadyakytlos et al., 2013).

Smith and colleagues (1988), for example, conducted a series of experiments on asset markets (like stocks) using students. Contrary to the predictions of neoclassical economics, the researchers found "bubbles," an artificial inflation of value. These asset bubbles persisted through repeated experiments, so Smith traveled to Chicago in 1989 and recruited experienced traders to take part in experiments identical to those he and coauthors had conducted with students. He believed that the traders would not bid up the prices to bubble levels as students had. However, he found that bubbles were no less severe using traders as subjects (Smith et. al., 1993).

The experience of Smith and his coauthors is not unique, as evidenced in this report's case studies. Fooks and colleagues (2016) found that students and landowners behaved quite similarly, with some differences (e.g., farmers tended to be less mindful of budget in the auction game, making bids even when budgets were low and the chance of receiving a payoff from bidding was small).

A more relevant question for most policy research is whether students provide a good test case for how a particular result will apply outside the laboratory. When specific institutional knowledge is an important feature of the market, students are less likely to be an appropriate participant pool. Results from field experiments may be less likely to agree with those from laboratory experiments in these cases (Levitt and List, 2005; Levitt and List, 2007).

A Cautionary Note

Experiments should not be thought of as substitutes for other methods of research, but rather as complements. As Deaton (2010, 2016) emphasizes, economic theory must play a key role in shaping experimental design in order to impart lessons. Barrett and Carter (2010) emphasize that blind faith in experimental methods can distort the research agenda and lead to biased judgment. The same concerns that have led researchers to embrace experimentalism—chiefly a concern that unobservable data can often drive results—are relevant when designing an experimental study. That is, the simple act of randomizing does not in and of itself mean that causality has been established. Some phenomena are difficult to study experimentally because, for example, they are geographically or institutionally (such as a village or State) specific, making randomization over similar entities difficult.

Even in cases when an experiment is scientifically desirable, there can be unique ethical considerations. Unlike observational studies, experimentation involves *intervention* into the lives of humans. The principle of equipoise suggests that experimental research should be conducted when there is genuine uncertainty about which treatment will benefit participants, implying that treatments known to benefit individuals should not be withheld simply for the sake of research. Furthermore, researchers should balance the interests of learning with the interests of study participants—even individuals not involved in a study who might still be affected (DeMartino and McCloskey, 2014).

Overall, experiments are most informative when researchers (and the policymakers who consult them) think critically about the questions that a study is designed to answer—when the research question determines the method of inquiry and not the reverse.

Case Studies: Policy-Relevant Experiments and the USDA

Here we review five policy experiments—three of which examine whether it is possible to improve USDA land conservation programs such as USDA's Conservation Reserve Program (CRP). Our primary goal is to illustrate the potential benefits, and relevant issues, related to the use of experiments by USDA agencies. Although the following studies focus on conservation and farm program administration issues, the USDA also has an active research program applying experimental methods to nutritional programs and policies

In each section, we describe the policy question and the experimental approach. We discuss how the experiment considered four experimental *design features*: control, context, internal validity, and external validity. Findings from the experiments are summarized and conclusions are noted, including how the results might influence policy.

A *Laboratory* Experiment: Asymmetric Auctions and the Design of Enrollment Mechanisms

Summary

Auctions are a cost-effective and efficient way to choose which potential participants to enroll in an oversubscribed program. There are many ways to run an auction, and that choice can lead to very different outcomes. Unique aspects of an auction's purpose and the potential participants means some designs will work much better than others. In this section, we examine a series of experiments that examined two different auction formats.

Researchers ran a series of lab experiments that used students as subjects. Although abstracted from real-world conditions, these experiments were customized to mimic some key aspects of conservation auctions, including the use of maximum values for bids ("bid caps"), variations in the amounts participants are willing to accept ("reservation prices"), and limits on the number of accepted offers (rather than a budget limit). The study revealed that when using an auction that placed "quota" limits on the number of accepted offers from similar bidders, enrollment goals could be achieved at lower cost than when using an "open" auction (where the number of offers accepted from similar bidders is not limited).

The Study

In a reverse auction,[17] all potential participants assign a unique value to the item being auctioned. This is their "reservation price"—it is their absolute "bottom-dollar" willingness to make a deal (the lowest amount they are willing to accept). Although participants in an auction know their self-assigned values, the buyer (such as a land conservation program) may be unable to differentiate between bidders before the auction begins, as the land associated with any given bidder has unique characteristics. Auctions where the buyer sees bidders as similar to one another are known as *symmetric auctions.*

However, in conservation auctions, the buyer actually might be able to observe important differences between bidders, a feature of reality (the "target environment") that a simple symmetric model does not incorporate. For example, the average productivity of land differs across the country, implying that farmers from more productive regions have higher earnings and therefore a buyer can observe indicators of their higher "bottom dollars."[18] As such, a conservation program administrator might be able to differentiate among bidders before the auction begins, resulting in an *asymmetric* auction, where participants are, measurably, not all the same.

When sellers are asymmetric, the most cost-effective auction treats different sellers differently (Myerson, 1981; Bulow and Roberts, 1989). How, then, might an auctioneer—who knows that the land that sellers are interested in enrolling in a conservation program are observably different—structure an auction that accounts for these differences?

[17]The term "reverse" auction simply refers to an auction in which a single buyer purchases from multiple sellers, as opposed to an auction in which multiple potential buyers compete for the right to purchase a good from a single seller.

[18]For example, since the typical acre of farmland in Iowa produces greater agricultural revenue than the typical acre in Texas, the typical farmer in Iowa will require a higher per-acre payment to enroll in a land conservation program than the typical farmer in Texas.

Researchers conducted lab experiments in a context where the buyer (an auctioneer) takes advantage of observable differences. In particular, the researchers tested if quotas on the number of accepted offers would increase competition[19]. In an auction with quotas, the auctioneer specifies *before the auction begins* that no more than a certain number of offers from defined classes of "similar" sellers will be accepted.

Each of these experimental auctions devised two "classes" of sellers—a "low opportunity cost" class and a "high opportunity cost" class (who required higher payments in order to sell to the auctioneer). Lab participants knew their own costs, they knew which class they belonged to, and they knew the range of costs within each class. And, the auctioneer knew which class each participant belonged to. These features are meant to mirror target environments (such as USDA's CRP), where participation is voluntary, potential participants may come from any region, farmers know their own costs, farms within a region tend to have similar agricultural productivity, and the program administrator has some information on each potential participant.[20]

The experiments included both a quota auction and an open auction. The quota and open auctions were identical except that the open-auction buyer could decide to enroll every bidder in the low-opportunity-cost class before enrolling any high-opportunity-cost bidders. Because of this, low-opportunity-cost bidders—those with the greatest willingness to make a deal—have less incentive to bid low in an open auction. This can erode the cost-effectiveness of the auction.

Experimental design features

- <u>Control</u>

  In this experiment, the researchers knew the opportunity costs of the participants. In actual circumstances, the buyer (for example, in the CRP, the buyer would be USDA's Farm Service Agency) usually has imperfect knowledge of these opportunity costs. The researchers wanted to test whether using a quota auction alters the bidding and improves outcomes within the asymmetric auction format.

  Because the researchers supplied the opportunity costs for each participant, the experiments could determine when behavior differed systematically across treatments. Each session included two treatments: open auction and quota. Because everyone participated in both treatments, the researchers could compare how the same individual performed under both types of auctions, and compare the relative performance of different participants. In short, differences in the observed bidding behavior were likely due to the different auction formats rather than other factors.

[19]This experiment and its findings are described in detail in the ERS report, *Options for Improving Conservation Programs: Evidence from Economic Experiments on the Design of Enrollment Mechanisms* (Hellerstein et al., 2015).

[20]For example, the CRP conducts an asymmetric auction that uses specific information about each parcel to select offers for general sign-up enrollment (USDS/FSA 2016b). This information includes local rental rates, which are used to determine each offer's Soil Rental Rate, and information on the characteristics of a parcel of land. The Soil Rental Rate is used to set the maximum acceptable bid (bid cap). The land characteristics, along with the conservation practice the landowner will install and the payment requested by the landowner, are used to create an Environmental Benefits Index for each parcel that is used to rank offers (USDA/FSA 2012).

- Context

  Researchers investigated quota auctions in a series of experimental sessions using undergraduates. The participants bought and sold "tickets" rather than conservation contracts to avoid bidding behavior that might stem from pro-conservation or anti-conservation preferences.[21] Each ticket was assigned an opportunity cost (in either the low- or high-cost category), so that any ticket purchased by the buyer earned the student the asking price minus the assigned cost. Participants also knew the cost ranges within both category of tickets. Furthermore, in the open auction treatments, each ticket had a bid cap—the participant could not request an amount greater than the bid cap.

  Participants were incentivized to perform well in the auction. Participants who "won" the auction (whose bids were accepted) collected a cash payment, so that an individual's actions could lead to greater take-home earnings.[22]

- Internal Validity

  Because researchers knew the underlying opportunity costs, which students earned if their "tickets" were not accepted, they could confidently determine the impact of each auction type on the change in behavior and overall outcomes.

  One threat to internal validity in this study, and others like it, is that participants behave differently at the beginning of an experiment, with limited experience in bidding, than they would having acquired experience. If, for instance, the single quota auction was always administered first, followed by a single open auction, quota auctions might appear to be more or less cost-effective simply because participants are initially more naïve in their bidding behavior. To combat this, researchers used practice rounds to ensure that participants understood the rules of the auction and received feedback before any consequential auctions were held under either auction design.

- External validity

  Since this experiment dealt with an auction design that is a significant departure from existing policy, its designers did not seek to maximize external validity. Thus, the external validity of these results is uncertain. First, as with most lab experiments (and some field experiments), the participants in this study were volunteers who had previously expressed interest in participating in economic experiments. Students who choose to participate are not necessarily representative of all students at a university and, more importantly, are unlikely to be demographically representative of participants in the target population. Will farm landowners in an actual quota-based auction behave in the same way as these students? For example, farmers might optimize their bidding behavior in a conservation program (such as the CRP) differently than students. Also, in an actual program, important tradeoffs may exist between encouraging lower bids and obtaining greater environmental benefits given that offered parcels can vary both in prices and in the level of environmental characteristics—a feature that other auction designs can account for (Hellerstein et al., 2015).

---

[21] In the technical language of experimental economics, the experimenters used "neutral language."

[22] A cash payment used to incentivize participants is a common practice in experimental economics (Smith, 1976; and Smith, 1982). It is, however, possible to conduct hypothetical experiments in which cash payments are not used. These experiments are less expensive, all else equal, and can be a valuable tool (Bardsley et al., 2010).

Other differences between student participants in the experimental environment and landowner participants in the target environment include frequency of bidding. During a laboratory experiment, participants are able to bid in many auctions over a relatively short period of time; in a policy environment, auctions can take place infrequently.[23] The ability to translate bidding strategy (how much to bid) into results—how much a participant earns—might be an important driver of the results.

Results

In both standard and quota auctions, participants with higher opportunity costs (more productive land) reduced their bids to approximate their opportunity costs. In the standard auction, participants with low opportunity costs were better able to increase their bids above their opportunity costs (leading to higher costs to the buyer). The quota auction, on the other hand, caused bidders with lower opportunity costs to bid much less than in the standard auction. Thus, the quota made low-opportunity-cost sellers act as if there were increased competition. The competition caused them to bid up to 14 percent lower, enhancing the potential cost-effectiveness of the auction.

Broadly speaking, this experiment revealed two insights. First, if the differences in opportunity cost between groups are large, substantial savings can be achieved by imposing quotas that limit the number of offers accepted from each group. Second, if between-group differences are minor, total procurement costs do not increase much through the imposition of a quota—a quota is a relatively "safe" mechanism from the perspective of the buyer. This does not mean that using a quota will always reduce the cost to the Government of carrying out competitive enrollment. Total costs could increase if (a) the quota leads to acceptance of a high-cost bid instead of a lower cost bid that is rejected due to a quota, or (b) members of the low- opportunity-cost group do not reduce their bids.

Conclusions

The virtue of this type of laboratory experiment is the high degree of control, which provides high internal validity when testing treatments. In this case study, researchers were able to confidently conclude that using a quota auction mechanism rather than an open auction mechanism can reduce overall costs to the buyer (the Government).

Lab experiments offer a first test of policy-relevant mechanisms that might subsequently be tested outside the lab. The obvious weakness of a policy experiment in the lab is external validity. For policy mechanisms that survive lab testing, a useful next step in evaluation is a test that attempts to compare outcomes in a target environment more similar to the environment in which a policy would operate. One type of test that naturally follows a laboratory experiment is an **artefactual** *field experiment*—a lab experiment using "real world" participants (rather than easy-to-enroll students).

[23]For example, the CRP's "general signup" auctions take place on an irregular schedule, typically no more than once per year. "Continuous signup," however, is ongoing and does not involve auctions.

Comparing Student and Artefactual *Field* Experiment: Do Bonuses Increase Contiguous Enrollment in a Conservation Program?

Summary

Findings from many different disciplines suggest that the environmental benefits generated by conservation programs (such as the CRP) may depend on the spatial configuration of the enrolled parcels. If parcels are adjacent to one another rather than scattered throughout the landscape, the same number of acres may provide higher benefits. For example, wildlife habitat for certain species needs to be large enough to allow the animals to thrive; and longer contiguous stretches of buffered riparian areas lead to greater reduction in sediment and improvements in water quality.

Conservation programs usually do not provide incentives to landowners to address this objective—not many programs specifically select parcels to achieve contiguity. Because there are few real examples, using a laboratory experiment is one way to examine how different forms of such a program could operate and demonstrate likely outcomes. In this section, we present an *artefactual field experiment* (with farmer participants) that compared several design mechanisms to achieve adjacency. These experiments highlight the potential uses of strategic targeting and drawbacks when using network bonuses.

The Study

Researchers examined whether ***network bonuses***—payments made when enrolling a parcel near already-enrolled parcels—can enhance overall environmental benefits. The efficacy of ***spatial targeting***—where program administrators select to enroll parcels that are near already-enrolled parcels—was also examined. Both design options are meant to encourage enrollment of parcels that are adjacent to one another (Fooks et al., 2016).

Two sets of experiments were run: one with students and another with agricultural landowners. Both sets of study participants bid in an auction and faced opportunity costs. Participants were presented with scenarios where they owned several "agricultural fields of 100 acres each" that could be offered to a "conservation program."[24] However, in this case, network bonuses were paid to each field when an adjacent field enrolled. These bonuses created an incentive for neighbors to enroll, in addition to the preexisting incentives to make money by bidding above opportunity cost. The other mechanism available to the program managers was spatial targeting, favoring adjacent fields over nonadjacent fields.

Experimental Design Features

- <u>Control</u>

  In this experiment, the researchers wanted to determine what type of incentive or enrollment approach would result in the highest degree of social benefits for a given cost of enrollment.

[24]In contrast to case study 1, in this case study, the instructions and descriptions given to participants had an agricultural flavor—with fields having "agricultural returns" ranging from $200,000 to $800,000 per field. Participants could offer these for sale to a "conservation agency." Furthermore, while the budget varied between rounds, in any one round the overall budget of the "conservation agency" was set, whereas in case study 1 there was no budget constraint.

Would landowners work together to enroll adjacent fields in hopes of receiving bonus payments? Would landowners who worked together achieve higher payments, leaving fewer remaining dollars to enroll other fields? Or would landowners ask for less than their opportunity costs because the network bonuses would provide additional compensation?

To simplify the experimental design, all fields had the same stand-alone environmental benefits; thus, the only way to increase a field's value to the conservation program was through adjacency. Fields joining an existing network of adjacent fields had a higher probability of being accepted with spatial targeting and, in some cases, received a bonus if accepted. Information was provided to each participant about his or her opportunity cost, but as in the previous experiment, participants did not know the same information for other participants. Students could earn an average of $30, while landowners could earn an average of $75.[25] In some sessions, participants within the same "geographic" area could communicate with each other; in others, they were not able to communicate.

- Context

In this experiment, a participant could offer his or her field for enrollment in a conservation program and receive a payment if selected. If a field was not offered to the program, or if it was offered but not accepted, the participant received the field's "agricultural return." Participants each had three fields that they could offer to enroll, or retain in agricultural use, within one of three geographic areas. Adjacency was defined as having a common border longer than a point (i.e., not just a corner) within the same geographic area.

- Internal Validity

Because all attributes of each field are known, the researchers could assess whether bidding behavior changed due to the introduction of network bonuses. Similarly, the researchers could determine how well spatial targeting by the buyer (program administrator) functioned.

- External validity

Following the student experiments at the University of Delaware, the same experiments were moved into the "field." These artefactual experiments had lab-like control but used agricultural landowners instead of student participants to increase external validity. There are limitations, however. For example, in the target environment (say, a region with a new conservation program), landowners would have unknown agricultural income and their land unknown environmental benefits, both of which can affect bidding decisions and the relative value of the field to the program administrator. In this experiment, incidental adjacency—adjacency purely as a result of regular enrollment—was much more frequent than it would be in a larger environment, such as the enrollment area in many conservation programs or when the nationwide scope of the CRP is considered. This anomaly is undoubtedly a driver of the results reported below and an important caveat regarding the findings. In addition, the multiple rounds of the experiment give participants a chance to learn, whereas in the target environment enrollment opportunities may be infrequent.

[25]It is common in experimental economics to compensate participants at a rate that is consistent with their own opportunity cost of time. Since landowners have, on average, a higher opportunity cost of time than students, the landowner payments were concomitantly higher.

Results

In this experiment, spatial targeting increased total environmental benefits and social welfare. Surprisingly, network bonuses *decreased* total environmental benefits and social welfare relative to spatial targeting. The cost of the bonuses used much of the budget, resulting in fewer fields being enrolled. With network bonuses, landowners of high-benefit fields placed offers more often; spatial targeting, on the other hand, allowed the buyer (such as the USDA, Farm Service Agency) to select the best options from the high-benefit fields. Letting the participants communicate had little impact on the results.

In many respects, the students and the landowners behaved similarly. As to differences, the landowner experiment resulted in higher environmental benefits (40 percent higher) than those run with student participants. Landowners tended to pay less attention to the program's budget (which varied over the several periods of each auction experiment), i.e., they would make bids even when budgets were low. They also learned (in subsequent bidding) faster than the students.

When network bonuses were available, the experiments revealed that more fields near networks entered the auction. Therefore, the program administrator had more adjacent fields from which to select. The increased participation of adjacent fields, however, only increased environmental benefits when spatial targeting made the selection of adjacent fields more likely than nonadjacent fields. In particular, when the buyer offered bonuses without spatial targeting, the positive impact of having adjacent fields did not compensate for how the bonuses drained the budget (reducing the total number of fields enrolled in the program).

Conclusions

Further experimentation in the laboratory could explore a number of contiguity features, such as whether smaller bonus payments would still encourage participation without depleting the budget. Beyond this artefactual field experiment, a pilot program or randomized control trial could be conducted to ensure that the results have external validity. Following a successful field experiment, one could determine the degree of scalability in introducing the new program design to a broader set of potential landowners.

Importantly, although this experiment was clearly artificial—the experimental and target environments were markedly different—it did highlight an important tradeoff: the need to balance the incentive effects of network bonuses against the possibility that paying these bonuses will reduce the pool of payments available to enroll more parcels.

A Policy-Relevant *Field* Experiment: Personal Discount Rates

Summary

Farmers' discount rates—a measure of the amount of compensation farmers require when they agree to receive payments at a later time instead of today—play a large role in their decisions regarding crop rotations, input choices, investment in machinery or structures, adoption of new technologies, and provision of environmental services, including participation in conservation programs. Policymakers may be especially interested in discount rates because farmers make many longrun decisions that impact environmental quality. For example, payment plans that better match the temporal preferences of farmers can have lower multiyear costs to funders of conservation programs.

In this section, we present a *framed field experiment* (with farmer participants) that asks participants to decide between receiving payments now or in the future. When presented with a choice between an immediate payment of several hundred dollars or receiving more by waiting 9 months, the average discount rate for farmer participants was an estimated 34 percent. That is, farmers are predicted to accept 34 percent less for an immediate payment.

The Study

Researchers collaborated on a field experiment—one conducted outside the lab with nonstudent participants (Duquette et al., 2012). The experiment was designed to elicit the discount rates of farmers. A precise understanding of the discounting behavior of farmers can help improve the design of conservation programs when those programs engage farmers in long-term contracts or seek to incentivize structural practices that have long lifespans. With respect to long-term contracts, farmer discount rates influence whether it is more cost effective to increase enrollment through higher annual payments or higher incentive payments at signup.

For example, some USDA programs (such as the CRP) provide annual payments to farmers in exchange for conservation efforts; other programs (such as the Wetland Reserve Program) purchase easements using one-time cost shares or enrollment payments. While initially more expensive, the use of upfront payments could increase enrollment and reduce long-term costs to the Government by aligning the timing of payments with farmer preferences.[26]

Experimental Design Features

- Control

  Working in an experimental lab setting using participants drawn from the target population (landowners) allows researchers to know and control many of the characteristics that may affect participants' decisions. For this experiment, the researchers ceded some control in order to observe real farmers who were facing a real decision about their farms and payment structure. In this case, the experiment's designer knows little about each participant's current agricultural

[26]If the Government has a lower discount rate than farmers, payments upfront can lower the Government's long-term costs and encourage more farmers to participate. It should be noted that in many programs, the payment structure is statutorily defined, hence modifications of payment schemes may not be straightforward.

income and how the experimental payment will affect the farmers' well-being. Neither do they know the exact level of benefit society would derive from a farmer choosing one of the options. However, a field experiment allays concerns that farmers might make different choices in a lab environment than when faced with similar time-money tradeoffs in their everyday lives.

- Context

  To measure discount rates, researchers conducted a natural field experiment where farmers were offered a choice between a *Now* payment of roughly $400 and a higher *Later* payment that varied across participants.[27]

  The farmers chosen to receive these offers had not been previously informed that a payment would be forthcoming and therefore would not have expected them. The offer process had several features:

  1. Participants faced a single choice, much like other economic decisions they might make, so that they would not readily see themselves as part of an experiment;
  2. Participants' choices paid out real money, and the stakes were large in comparison to those often used in laboratory environments;
  3. A credible reason was provided for why the payments were being offered and why a choice of payment timing was included,[28] and
  4. The time horizon was consistent with the timing of producer decisions.

- Internal Validity

  Farmers were randomly assigned one of three payment treatments (table 1), with a roughly equal number of participants in each treatment.

Table 1
Three experimental treatments of now and later payments, interest rate, and number of participants

	Now payment	Later payment	Implied interest rate	Number of participants
Treatment 1: small later payment	$405	$430	8 percent	95
Treatment 2: medium later payment	$405	$463	18 percent	99
Treatment 3: large later payment	$405	$498	28 percent	99

Source: (Duquette, Higgins, and Horowitz, 2012).

[27]American Farmland Trust (AFT) conducted the experiment with the farmers with whom they had an established relationship. For those who chose Now, AFT sent payments approximately 14 days after the decision was made. For those choosing Later, AFT sent payments approximately 9 months later.

[28]Individuals received an environmental stewardship payment from AFT, a well-known organization focused on agri-environmental stewardship. "By being a good steward you are producing valuable environmental services that we can all enjoy…. You will be receiving an environmental stewardship payment (of about $400!) as a result of the practices you have already implemented on your land." When describing the Later payment option, the letter stated, "To help our Finance Division spread these payments out, we are offering you two options on the timing and amount of the check you will receive."

Each farmer saw only one of the payment pairs. This choice was used to compute a single bound on his or her discount rate. If a participant chose the *Now* payment, his or her individual discount rate was bounded below by the implied interest rate shown in Table 1; for example, a farmer choosing the *Later* payment when offered the medium payment (treatment 2) pair could have had a discount rate greater than 18 percent.[29] Alternatively, a farmer who accepted the immediate payment in a treatment 1 pair would have had an interest rate less than or equal to 8 percent.

By randomizing farmers to receive one of three payment pairs and fitting responses to a structural statistical model, the experimental data can be used to estimate an average discount rate for the population of farmers involved in the experiment. Because the payment was framed as a payment for actions that farmers were already taking—being good "stewards" of the land, for instance—the response of farmers is arguably more representative of the decisions they would make in their everyday lives than a decision logged in a laboratory environment. This was done to increase the likelihood that the experiment accurately measures discount rate, rather than other factors, such as experimenter demand effects[30] that might otherwise influence decisionmaking.

- External Validity

Farmers frequently make decisions that are influenced by their own personal discount rates. Therefore, a natural setting exists under which these rates can be measured via a field experiment. Within a natural field experiment, participants generate data in the course of their normal actions. For this reason, natural field experiments are conducted in such a way that participants do not know that they are part of an experiment (Harrison and List, 2004). This is desirable for at least two reasons: (1) participants will not change their behavior simply because they are being observed, and (2) participants will use their normal decisionmaking processes rather than processes that are particular to behavior in a lab. Regardless of how participants make economic decisions, be it rule of thumb or sophisticated calculus, a field experiment that does not disturb these processes generates data with arguably higher external validity.

External validity amounts to the generalizability of findings from an experiment to other environments. In this case, the experiment can be said to have high external validity within the population studied, though the measured average discount rate is not necessarily representative of all farmers in all circumstances. For instance, farmers who have dealings with the American Farmland Trust (AFT) may be different than farmers at large.[31] It is also possible that farmers make different time-tradeoff decisions about receiving direct financial payments than they would about something like investing in a new technology—a decision that involves actively making payments today in order to enjoy benefits later. It is also possible that discount rates are sensitive to the *scale* of payments, so that discounting is heavier, for instance, at lower payment levels than at higher, more realistic levels. Although this experiment used payments larger than

[29]If the participant chose a later payment, then his or her individual discount rate was bounded above by the corresponding interest rate.

[30]"Experimenter demand effects" refers to the fact that individuals might behave differently when they know that researchers are observing them and might particularly be influenced by what they think "appropriate" decisions might be. That is, when individuals are actively observed, they might make decisions based not just on their own preferences but also on the preferences of others in the experiment, including the researchers. This phenomenon, known as the Hawthorne Effect, states that participants in experiments and real life may change their behavior simply because they know they are being observed (Levitt and List, 2011).

[31]In fact, it is possible, as a reviewer noted, that individuals having dealings with AFT, and who support AFT's mission, might have chosen the smaller payment (the sooner payment) so as to save AFT money.

those typically offered in laboratory experiments, they were still lower than the stakes in many decisions that farmers make about their business.

Results

Out of 208 farmers who responded to the solicitation,[32] 147 (70.7 percent) chose the *Now* offer, which corresponds to an average discount rate of 34 percent across all respondents. Eighty-seven percent of farmers chose to receive the *Now* payment ($405) today when faced with the alternative of $430 in 9 months, indicating an annual discount rate of at least 8 percent (table 2). As the later payment amount increased to $463 and $498, fewer farmers chose *Now*, dropping to 66 percent and 60 percent, respectively. While this average seems quite high, similar results have been found in numerous experiments (Frederick et al., 2002). When regression analysis is used to control for both payment size and timing (table 3), the estimated discount rate is 34 percent.

The discount rates estimated in the AFT field experiment exceed the estimates using farm-level data from other studies (Abdulkadri and Langemeier, in 2000, reported an intertemporal substitution (which is akin to a discount rate) of 0.158 to 0.351; Lence, 2000). The AFT rates are closer to studies that directly elicit discount rates, such as Harrison and colleagues (2002), and that exploit natural experiments, such as Warner and Pleeter (0 to 30 percent) (2001).

Table 2
Number and share of farmers choosing payments now

Level	*Now* choices	Total choices	Share choosing *Now*
Small: $430	59	68	86.8%
Medium: $463	44	67	65.7%
Large: $498	44	73	60.3%
Total	**147**	**208**	**70.7%**

Source: Duquette, Eric, Nathaniel Higgins, and John Horowitz (2012). “Farmer Discount Rates: Experimental Evidence,” *American Journal of Agricultural Economics* 94(2): 451-56.

Table 3
Estimates of the mean discount-rate parameter for all farmers using a maximum likelihood procedure

Discount rate (δ)	0.34*** (7.60)
Observations	208
Log likelihood	-119.8

Notes: z-statistics are reported in parentheses and use robust standard errors; *** p<0.01

Source: Duquette, Eric, Nathaniel Higgins, and John Horowitz (2012). “Farmer Discount Rates: Experimental Evidence,” *American Journal of Agricultural Economics* 94(2): 451-56.

[32]While 293 farmers were sent the letter, only 208 farmers responded for a 70.9-percent response rate.

Conclusions

This experiment's results suggest that environmental programs such as the CRP could increase enrollment or reduce costs by changing the timing of payments to better reflect farmers' preferences. For example, a conservation program that leases land over a 10-year period could give farmers a lump-sum payment worth 10 years of annual rental rates upfront. This may attract participants with high discount rates who otherwise would not join. The conservation program[33] could offer farmers an option of a smaller lump-sum payment. Instead of 10 annual payments of $100 per acre ($1,000), the program could pay out $80 per acre in a lump sum ($800), saving the taxpayers 20 percent of the cost of enrollment. Based on AFT field experiments, many farmers would prefer a lump sum.

In contrast to the *policy experiment in the lab* or the *artefactual policy experiment with farmers, a policy experiment in the field* has enhanced external validity on some dimensions (as discussed above). The field experiment was a follow-on to experiments researchers had conducted among students in the lab. Testing the discount behavior of farmers in the field, though comparatively expensive, is required to verify that the high discount rates observed in the lab exist for farmers in the real world.

Still, however, this policy experiment in the field remains one step removed from a government program. The farmer participants made real decisions for real money, but did so in a unique context created by the researchers. A final step is to conduct an administrative experiment directly within the workings of a government program.

[33]The CRP is an example of a conservation program that leases land over a fixed number of years. Of course, the reasons for participation in programs like the CRP are myriad, including guaranteeing a fixed income stream and serving a conservation ethos. Thus, although these results may be suggestive of how participation in a program might change with payment timing, they are not definitive.

An *Administrative* Experiment: Encouraging Participation in the CRP

Summary

The Conservation Reserve Program (CRP) has been the largest conservation program in the United States for many years. The extent to which farmers are willing to participate is a crucial determinant of program outcomes, both in terms of environmental benefits and program costs. This section describes an administrative experiment where a large number of farmers eligible for the program are contacted with outreach letters. The study finds that the use of these "nudges" can increase participation rates, at a cost of around $41 per additional offer.

The Study

ERS partnered with USDA's Farm Service Agency (FSA) to examine potentially cost-effective ways of increasing CRP participation. Based on prior studies of water and energy conservation, ERS and FSA decided to investigate the impact of informational outreach on participation. The 2012 CRP General Signup was selected to test the effectiveness of outreach given the large percentage of CRP acreage with expiring contracts that year.

During the first week of the 4-week General Signup period, ERS and FSA randomly selected over 100,000 farms to receive a reminder letter about the signup (Wallander et al., 2015).[34] Given the program's high visibility and the fact that many of the letter's recipients had expiring contracts, this experiment involved a population already informed about the program. The experiment tested the behavioral economics hypothesis that individuals often ignore or do not respond to relevant information—a form of behavior that is often described as the result of bounded rationality. In addition to this hypothesis, the experiment tested if information such as peer comparisons and social norm references would influence behavior. Alternate versions of the letters contained framing "nudges"—such as noting how many farmers have already signed up in the State or explaining how the installation of native grasses can improve the environment—that have proven effective in energy and water conservation.[35]

Experimental Design Features

- Control

  Although FSA has varied its outreach efforts over the years, it is difficult to disentangle the effects of variation in outreach from other factors that influence participation and that change over time. For example, outreach efforts often vary across States or counties within a given year, as do farmers' expectations about their future returns (against which they are weighing the decision to participate in CRP). If outreach efforts are intensified in areas where non-CRP income expectations are increasing, comparison of those counties to lower outreach counties could understate the impact of outreach. The environmental benefits that would have accrued from land not in the program constitute another unobserved variable that could confound efforts to evaluate the

[34]These letters were in addition to other baseline outreach materials and efforts.

[35]An example letter is included as an appendix.

impact of outreach. Researchers generally know the physical characteristics (such as erodibility) of a given farmer's land only if that farmer makes an offer to the program. These unobserved differences in environmental characteristics can influence willingness to participate regardless of outreach intensity.

Given that the impact of outreach efforts on CRP participation is likely to be small relative to other factors (such as farmer attitudes about conservation), it is critical to design an experiment in which these other factors are balanced (distributed equally) across the treatment and control groups. Obtaining sufficient statistical robustness for this experiment, then, requires a fairly large population. Using existing administrative data minimizes the cost of obtaining a sufficiently large sample because one can examine the administrative data of enrollments to detect responses to the treatment.

- Context

Because the experiment was conducted with the program agency (FSA) through the program itself, the context matches the actual policy. Reality is embedded.

Defining the appropriate groups to include within an administrative experiment is an important step. Therefore, before conducting the experiment, FSA and ERS needed to define eligible landowners for the outreach effort, as not all farmland is eligible to participate in CRP.

  - All farms with expiring CRP contracts were automatically eligible to submit an offer to re-enroll in the program. In addition, the farmers who own these fields are familiar with the program, and FSA had contact information. The researchers identified about 47,000 farm operations with contracts expiring in 2012. FSA had already sent letters letting them know that their contracts were set to expire.

  - Farms with eligible but unenrolled fields were more difficult to identify. Prior studies indicate that about 212 million acres meet CRP's eligibility requirements. Since eligibility is determined based on field characteristics such as erodibility and crop history, ERS/FSA used geographic information system (GIS) data to screen more than 30 million fields and identified over 900,000 farm operations with potentially eligible but unenrolled land.

- Internal Validity

The internal validity of the experiment depends on the assumption that random assignment to the treatment (letters) group balances unobservable factors that may influence the decision to participate in the CRP. We balanced the experimental sampling over observable variables—primarily the location (State) of the parcel—that are likely to influence participation. Uncontrolled factors that may influence participation include more local effects such as the county in which a parcel lies (since enrollment requires interaction with a local official, who might vary from county to county) and the profitability of land (the opportunity cost of participation). By using random assignment across a very large sample size, the experiment's design seeks to balance these uncontrolled factors between the treatment and control groups.

• External Validity

By implementing the field experiment within the actual CRP program, this experiment bypassed the external validity concerns of many other experimental designs. If a statistical relationship were found between receiving the letter and/or individual messaging (peer comparisons and social norm references) and re-enrollment, this could be scaled up to the entire group of eligible farmers. As long as the population involved in the experiment is representative of the entire population of potential CRP participants, and the conditions of CRP do not change markedly, the effect of messaging estimated in this experiment is likely to persist.

Results

Outreach letters sent during the first week of the signup period resulted in an additional 16 to 17 offers per 1,000 reminders sent to farms with expiring contracts. This translates into a printing and mailing cost of $41 per additional offer.[36] In this context, a standard letter, one with peer comparisons, and one with social norm references were each equally effective as outreach reminders. The peer comparisons (e.g., "Did you know that 3,240 farmers in your State have already signed up?") and social norm references ("The most competitive contracts received 65 or more additional EBI (Environmental Benefit Index) points for using improved practices like planting native grasses. These stewards have helped improve water quality and wildlife habitat for their neighbors and protected farmland for future generations of farmers. Everyone wins with high EBIs.") that have been effective in other contexts did not affect participation differently than a standard letter. The main effect came from reminding farmers about the opportunity to participate rather than from the particular content of the letter.

For potentially eligible but unenrolled farms, the letters did not boost enrollment. One reason is that very few parcels within the control group of eligible and unenrolled lands are offered for enrollment. A treatment effect of 1.68 percent, as found for those with expiring acreage, would represent an increase of over 800 percent in offer rates for the population with eligible but unenrolled land. Part of this fixedness may relate to the quality of data (names and addresses) on farms not currently participating in the program, as well as limited ability to link the data to offers.

Conclusions

Increasing the number of land parcels offered can result in cost savings and higher environmental benefits through the CRP's competitive bidding process. This simple experiment enabled ERS to place a dollar value on USDA's primary mechanism to generate additional offers for its flagship conservation program. USDA can use this information to better use outreach funds to encourage participation. Extrapolated beyond the 2012 signup, the decision to stop sending out reminder letters would reduce offers by an estimated 500,000 acres for a 25-million-acre program over a full 10-year cycle.

[36]The timing of the letters—sent during a signup period—may have influenced their effectiveness. However, the experiment did not vary timing, hence the experiment's results cannot be used to test this possibility.

An Administrative Experiment: Encouraging Voter Participation in FSA County Committee Elections

Summary

USDA's Farm Service Agency interacts directly with farmers and ranchers through a network of local field offices, where farmers can inquire about or apply for programs. The management of these field offices is overseen by a committee whose members are elected by all farmers eligible to participate in FSA programs. Over time, participation in County Committee (COC) elections has declined. This section examines an administrative experiment in which nearly all farmers eligible to vote were contacted with specially designed ballot mailers and outreach letters/postcards. The experiment finds that information in the letters and postcards can increase participation by several percentage points at a cost of less than $29 per potential voter.

The Study

USDA's Farm Service Agency (FSA) operates programs—such as loans, disaster payments, and commodity/conservation programs—that affect the lives of farmers and ranchers, their income, and the economy. Each field office is administered by a County Executive Director who is responsible for the local implementation of FSA programs. The County Executive Director is, in turn, overseen by the county committee. In addition to being a point of contact between FSA and farmers, county committees are critical to the delivery of farm support programs and make numerous program determinations on issues such as whether or not a producer is in compliance with program eligibility requirements.

County committees consist of three to five members, depending on the number of local administrative areas (LAAs) that the committee represents.[37] A single elected committee member represents each LAA for a 3-year term.[38] COC members must live in the LAA that they represent.

County committees were first authorized by the Soil Conservation and Domestic Allotment Act of 1935. Over time, participation in COC elections has declined. In an effort to increase participation in vote-by-mail balloting, FSA partnered with ERS and the National Science and Technology Council's Social and Behavioral Sciences Team (SBST) to test changes to COC election ballots and outreach materials.

Experimental Design Features

The experiment was conducted during the 2015 COC elections, which took place by mail. FSA mailed a ballot to each eligible voter in early November, with a deadline for returning it approximately 1 month later. Two additions to voter mailings were tested in the experiment: (1) candidate information printed on the outside of ballots, and (2) postcards with candidate information.

One barrier to casting a valid ballot may be the simple act of opening the envelope and evaluating candidate choices. This barrier might be especially strong if previous experience with the COC elec-

[37]The name "county committee" is somewhat of a misnomer as the committees frequently oversee operations in areas that cross county boundaries.

[38]Each committee member can serve for a maximum of three consecutive 3-year terms.

tion process has led the eligible voter to downgrade the value of voting. This might occur because COC elections are often uncontested, with just one candidate running for the open seat, a fact that eligible voters may consider when they receive a ballot in the mail.

COC elections are often characterized by limited public information about candidates, and differentiation between candidates is generally low. In such elections, providing easily viewed and compelling information about the candidates might help overcome the simple barrier to opening and submitting the ballot. To test this hypothesis, researchers printed the names of candidates—which are otherwise included only on the inside of sealed ballots—on the outside of some ballots so that they would be readily apparent to eligible voters.

Another barrier to robust election returns is that voters may simply forget to vote by the deadline, even if they intend to. To test this effect, two postcards were sent to all eligible voters who were assigned to the relevant treatment condition, one designed to arrive approximately 1 week before the ballot arrived in the mail, and one designed to arrive approximately 1 week before the submission deadline.

The pre-ballot postcard included (1) the names of all candidates running for election; (2) a personalized message encouraging eligible voters to help make sure farmers in their county were represented; (3) a reminder that the term of the committee member would be 3 years, implying that the next chance to vote would be 3 years hence; and (4) a picture of the ballot that would be shortly arriving in the mail.

The post-ballot postcard included the same information as the pre-ballot postcard, with an additional reminder that the deadline for submission was approaching. Voters were also informed that if the ballot had been lost (or failed to arrive in the mail), the eligible voter could obtain a new ballot by visiting that person's local field office.

- <u>Control</u>

  To test the effects of both the on-ballot candidate-information treatment and the postcards, each eligible voter was assigned to receive one of four treatments (table 4).

Table 4
Treatments in county elections study

	Postcards	No Postcards
Ballot with candidate names on outside	T1	T3
Ballot without candidate names on outside	T2	T4

Source: USDA, Economic Resesarch Service summarization of data design.

Eligible voters assigned to treatment one (T1) received both pre- and post-ballot postcards, as well as a ballot with candidate names printed on the outside. Those assigned to treatment two (T2) received the pre- and post-ballot postcards and a standard ballot. Those assigned to treatment three (T3) received a ballot with candidate names printed on the outside but no postcards. Those assigned to treatment four (T4) received a standard ballot and did not receive postcards. All groups received the standard mailings from FSA.

The assignment of voters to the four treatments was done at the "household" level,[39] such that all ballots sent to the same address, regardless of recipient, received the same treatment. This was meant to prevent confusion that might arise from different combinations of postcards and ballots arriving at the same address.

- Context

Because the experiment is conducted in conjunction with the program agency and is embedded in the program itself, the experimental context is identical to that of the target environment. Reality is an inherent feature of the design.

In addition, the entire eligible population of farmers is included in this experiment.[40] Each eligible voter is identified before the election by FSA using administrative data, and each eligible voter is sent a ballot. Some additional voters may enter the population if they do not receive a ballot—but believe that they are eligible—by applying for a ballot at their local field office. The number of voters who do this is small, however, and they are excluded from the analysis since they were not assigned to any treatment group.

- Internal Validity

The internal validity of the experiment depends on the assumption that the random assignment of treatments effectively balances unobservable factors that may influence the decision to cast a vote in the COC election. We balanced the experimental sampling over observable variables that are likely to influence participation, including the election (which is specific to a local area) and the type of voter (an entity or an individual). All households containing a single type of voter (e.g., individual voters) within a single election have an equal probability of being assigned to each of the four treatments.

Because the treatments are assigned in a way that balances assignment locally across many individual elections, factors such as local get-out-the-vote efforts—which might vary from county to county—are unlikely to influence the estimate of average treatment effect since any such effort would affect households in each treatment.

As in every experiment, however, there are uncontrolled factors that could vary across treatment groups. For instance, individual voting histories—whether or not a given voter has voted in the past—are often found to be predictive of voting behavior in the future. Without historical voting data, researchers could not take this factor into account when assigning treatment groups. Consequently, individuals who voted frequently in past elections may have been unintentionally

[39]A household was defined as a single address, so that multiple voters receiving mail at the same address would be considered a single "household."

[40]Most experiments are conducted on a subset of the relevant population.

(and unobservably) assigned more often to treatment one, whereas individuals who have never voted may have been assigned more frequently to treatment four. By using randomizing treatments across a very large sample size of nearly 2 million voters across 1.4 million households, the experiment's design is likely to balance these uncontrolled factors across treatment groups.

- External Validity

  By implementing the field experiment within an actual COC election, this study overcomes the external validity concerns of many other experimental designs. Also, the entire eligible population was included in the experiment, allaying concerns that the sample is in some way selected, or different from the remainder of the relevant population. In this way, the experiment provides a precise measure of the effect of ballot treatments on the relevant population for the 2015 COC elections.

  Another external validity concern is whether or not the treatment effect would persist if repeated in subsequent years. If, for example, the environment of the 2015 elections was especially novel, a policymaker might doubt that the response rates would be similar if the experiment were repeated in 2016. The researchers are aware of no such differences.

Results

The turnout rate (rate of valid ballots submitted) for voters assigned to treatment four—those who received a ballot without information printed on the outside and who did not receive postcard mailings—serves here as the baseline rate of voter turnout. Existing administrative data indicate that the actual turnout rate among households in this group (i.e., at least one member of the household submitted a valid ballot) was 9.3 percent.[41]

We ran a multilevel model to estimate the effects of the treatments on voting behavior. The effects of each of the treatments—printing information on the outside of ballots and/or sending postcards—are estimated to be different from zero. The effect of sending the postcards is larger than the effect of simply printing information on the ballot (figure 4).

Conclusions

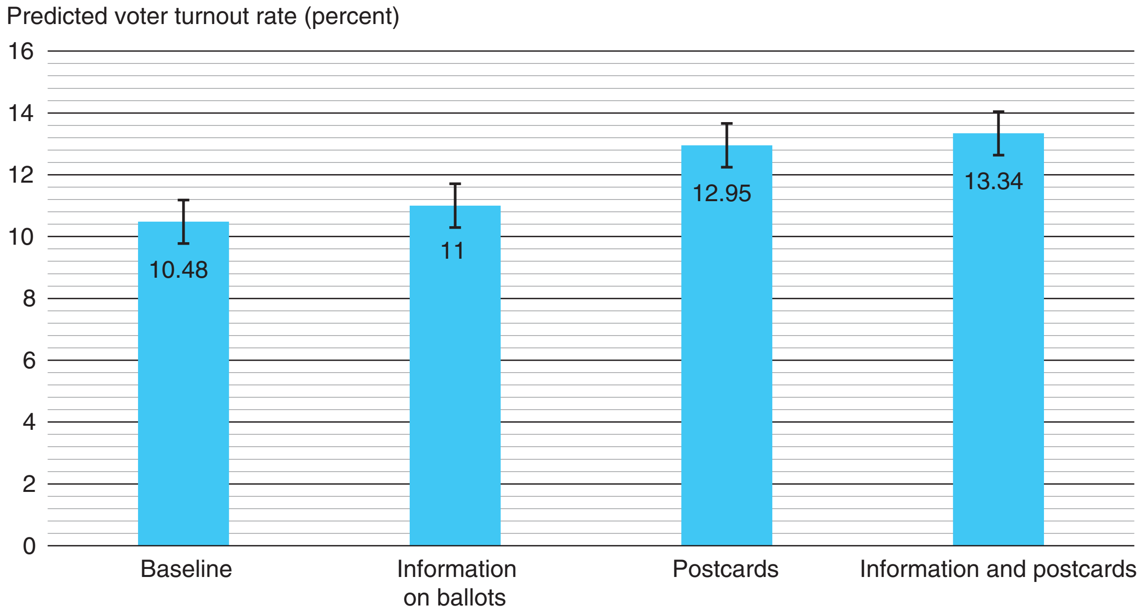
Increasing participation in the COC election process is important to ensuring local representation of FSA farmers in its programs. This simple experiment enabled precise estimates of the impact of simple changes in the way that election materials are presented to eligible voters. With a treatment effect of 2.9 percent (the increase in response rates between treatment T1 (postcard and information on the outside of the ballot and the baseline T4 (no postcard and no information)) and a postcard cost of $0.83 per unit,[42] this translates to one extra vote for every $28.55 spent. FSA can use this information to encourage participation in future elections and can build on the results here to create new low-cost outreach strategies.

---

[41]See Technical Appendix for details of the estimation.

[42]About 800,000 individuals received postcards, with the cost of each postcard made up of about 1.5 cents for printing, and about 40 cents for postage. Since two postcards were sent to each individual in the relevant treatment conditions, we round up to a cost of $0.83 per "unit," or individual.

Figure 4
Predicted turnout rate under different treatments



Notes: Vertical bars indicate confidence intervals. The text lists a 9.3% *actual* participation rate for the baseline, while the *predicted* rate listed above is 10.48%.
Source: County committee experiment data.

Conclusion

Economic experiments can be used effectively to improve the design and implementation of government policy and programs. Experiments are especially useful when evaluating new and innovative policies that are entirely *prospective*, since traditional economic techniques (such as statistical analysis of data and simulation modeling) may be unavailable or inapplicable for policies that have never before been implemented.

A broad spectrum of experiments exists that can be tailored to the questions being asked, the funding available, and the circumstances in which a policy will be implemented. Many occur in an academic laboratory using students responding to generic situations. Others have farmers making choices within programs. In many cases, because the researcher seeks control over the parameters, experiments do not completely reflect the actual environment of interest—experiments often take place in an environment that is meant to be a facsimile of a real economic environment.

Thus, when designing an experimental protocol, researchers need to consider the tradeoffs between having a high degree of control and the context of the experiment. In particular, does the experiment properly measure a causal effect of interest to policymakers given the sample population (known as internal validity)? And is it likely that these findings accurately reflect what would happen in the *target environment*—the real economic environment that the experiment is designed to represent (external validity)?

However, it can be difficult (and expensive) to achieve both goals. Focused experiments in a lab environment facilitate *control* and a high degree of internal validity, but this focus can reduce the generalizability. Such experiments may not be representative of the policy environment (the setting a policy would occur in or the people it would affect). Conversely, a field experiment using a cohort similar to the policy environment may be highly representative, but could have results affected by extraneous factors that are not relevant and are beyond the control of the experiment designer, requiring, for example, larger sample sizes to balance out factors that might confound estimation of the treatment effect.

Experiments can be especially productive when new programs or actions are being designed, when an existing program is oversubscribed and rationing is necessary, and when an existing program has inadequate enrollment. Setting up a design and/or an approach to evaluate a policy change prior to its rollout can improve the analysis, although it is not always strictly necessary. When administrative data are available, the cost and burden of conducting field experiments are reduced since the results of a treatment can be readily determined by examination of the "before" and "after" administrative data.

Economists have conducted experimental research on a variety of issues related to agricultural policy. This report surveyed five different examples of experiments that examined the possibilities of new auction designs in conservation programs; the use of agglomeration bonuses and targeting to achieve contiguity of land parcels; the discovery of farmers' discount rates; and the use of nudges and other behavioral mechanisms to encourage farmer participation in (1) the CRP and (2) FSA elections (table 5). These case studies illustrate how experiments can be applied and the types of insights they can provide.

Table 5

Characteristics of the five case studies

	CRP Quota	Contiguity	Discount rates	CRP nudge	County elections
Participants	Students	Students and farmers	Farmers	Sample of farmers	All eligible farms
Type	Lab	Lab	Field	Administrative	Administrative
Cost	Low	Medium	High	Low	Low
Experimental design features:					
Control	High—experimenter knew all details (values were "induced")	High—experimenter knew all details (values were "induced")	Medium—only choice set was under experimenter control; farmers had their homegrown (own personal) values	Low—little is known about other factors that may influence choices	Low—little is known about other factors that may influence choices
Context	Low—generic language used	Medium—agricultural example used	Medium high—offer process similar to real life, but consequences of choice are moderate	High—decisions are directly relevant	High—decisions are directly relevant
Internal validity	High—though repetitive nature could bias results through learning	High—though repetitive nature could bias results	Medium—randomization and modeling used to control for unobservables	Medium—depends on randomization over a large sample	Medium—depends on randomization over a large sample
External validity	Low—students in lab might differ from farmers making choices	Low to moderate—artificial nature of problem even though agricultural context used. Farmer versions more likely to be generalizable than those done with students.	Medium high—a sample of farmers facing a realistic choice	High—population in study pulled from the population of interest	Very high—population in study is the population of interest

Source: USDA, Economic Research Service.

References

Abdulkadri, Abdullahi O., and Michael R. Langemeier (2000). "Using Farm Consumption Data To Estimate the Intertemporal Elasticity of Substitution and Relative Risk Aversion Coefficients," *Agricultural Finance Review* 60(1): 61-70.

Ausubel, Lawrence, Peter Cramton, R. Preston McAfee, and John McMillan (1997). "Synergies in Wireless Telephony: Evidence from the Broadband PCS Auctions," *Journal of Economics and Management Strategy* 6(3), pp. 497-527.

Baicker, Katherine, Sarah Taubman, Heidi Allen, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Eric Schneider, Bill Wright, Alan Zaslavsky, Amy Finkelstein, and the Oregon Health Study Group (2013). "The Oregon Experiment—Effects of Medicaid on Clinical Outcomes," *New England Journal of Medicine* 368(18): 1713-1722.

Ball, Michael O., Lawrence M. Ausubel, Frank Berardino, Peter Cramton, George Donohue, Mark Hansen, and Karla Hoffman (2007). "Market-Based Alternatives for Managing Congestion at New York's LaGuardia Airport," in *Optimal Use of Scarce Airport Capacity, Proceedings of AirNeth Annual Conference,* The Hague, April.

Banerjee, Abhijit, and Esther Duflo (2008). "The Experimental Approach to Development Economics," National Bureau of Economic Research Working Paper No. 14467.

Barberis, Nicholas C. (2013). "Thirty Years of Prospect Theory in Economics: A Review and Assessment," *Journal of Economic Perspectives* 27(1): 173-96.

Bardsley, Nicholas, Robin Cubitt, Graham Loomes, Peter Moffatt, Chris Starmer, and Robert Sugden (2010). *Experimental Economics: Rethinking the Rules.* Princeton, NJ: Princeton University Press.

Barrett, Christopher B., and Michael R. Carter (2010). "The Power and Pitfalls of Experiments in Development Economics: Some Non-Random Reflections," *Applied Economic Perspectives and Policy* 32*(*4): 515-48.

Belot, Michele, Raymond Duch, and Luis Miller (2015). "Who should be called to the lab? A comprehensive comparison of students and non-students in classic experimental games," *Journal of Economic Behavior & Organization* 113: 26-33

Bulow, Jeremy, and John Roberts (1989). "The Simple Economics of Optimal Auctions," *Journal of Political Economy* 97(5): 1060-90.

Camerer, Colin F. (1989). "An experimental test of several generalized utility theories," *Journal of Risk and Uncertainty* 2(1): 61-104.

Campbell, Donald T. (1957). "Factors Relevant to the Validity of Experiments in Social Settings," *Psychological Bulletin,* 54(4): 297-312.

Campbell, Donald T. (1969): "Reforms as Experiments," *American Psychologist,* 24(4): 409-29.

Chamberlin, Edward H. (1948). “An Experimental Imperfect Market,” *Journal of Political Economy,* 56(2): 95-108.

Cody, Scott, Irma Perez-Johnson, and Kristen Joyce (2015). “Administrative Experiments: Unlocking What Works Better and What Works for Whom,” *American Journal of Evaluation* 36(4): 547-57.

Cox, James C., Vernon L. Smith, and James M. Walker (1988). “Theory and Individual Behavior of First-Price Auctions,” *Journal of Risk and Uncertainty* 1(1): 61-99.

Council for Agricultural Science and Technology (CAST) (2015). *Process Labeling of Food: Consumer Behavior, the Agricultural Sector, and Policy Recommendations.* Issue Paper 56. CAST, Ames, IA.

Cramton, Peter (1997), “The FCC Spectrum Auctions: An Early Assessment,” *Journal of Economics and Management Strategy* 6: 431–95.

Deaton, Angus (2010). “Instruments, Randomization, and Learning about Development,” *Journal of Economic Literature* 48 (2): 424-55.

Deaton, Angus, and Nancy Cartwright (2016). “Understanding and Misunderstanding Randomized Controlled Trials.” NBER Working Paper 22595, National Bureau of Economic Research, Washington, DC.

DeMartino, George, and Deirdre McCloskey, eds. (2014). *The Oxford Handbook of Professional Economic Ethics,* 1st ed., Oxford University Press.

Duquette, Eric, Nathaniel Higgins, and John Horowitz (2012). “Farmer Discount Rates: Experimental Evidence,” *American Journal of Agricultural Economics,* 94(2): 451-56.

Eden, T., and R.A. Fisher (1927). “Studies in Crop Variation: IV. The Experimental Determination of the Value of Top Dressings with Cereals,” *The Journal of Agricultural Science* 17(4): 548-62.

Eden, T., and R.A. Fisher (1929). “Studies in Crop Variation: VI. Experiments on the Response of the Potato to Potash and Nitrogen,” *The Journal of Agricultural Science* 19(2): 201-13.

Ellis, Sean F., Jacob R. Fooks, Kent D. Messer, and Matthew J. Miller (2016). “The Effects of Climate Change Information on Charitable Giving for Water Quality Protection: A Field Experiment,” *Agricultural and Resource Economics Review* 45(2): 319-37.

Exadaktylos, Filippos, Antonio M. Espina, and Pablo Bianes-Garza (2013). “Experimental subjects are not different,” *Scientific Reports* 3, Article Number 1213.

Filiz-Ozbay, Emel, and Erkut Y. Ozbay (2007). “Auctions with Anticipated Regret: Theory and Experiment,” *American Economic Review* 97(4): 1407-18.

Fisher, Ronald A. (1921). “Studies in Crop Variation. I. An Examination of the Yield of Dressed Grain from Broadbalk,” *The Journal of Agricultural Science* 11(2): 107-35.

Fisher, Ronald A., and W.A. Mackenzie (1923). “Studies in Crop Variation. II. The Manurial Response of Different Potato Varieties,” *The Journal of Agricultural Science* 13(3): 311-20.

Fisher, Ronald A. (1925). *Statistical Methods for Research Workers*. Genesis Publishing Pvt Ltd.

Fisher, Ronald A. (1935). *The Design of Experiments*. Oliver and Boyd, Edinburgh.

Fooks, Jacob, Nathaniel Higgins, Kent D. Messer, Joshua Duke, Daniel Hellerstein, and Lori Lynch (2016). "Conserving Spatially Explicit Benefits in Ecosystem Service Markets: Will Network Bonuses or Spatial Targeting be Effective?" *American Journal of Agricultural Economics* 98(2): 468-88.

Fréchette, Guillaume R. (2015). "Laboratory Experiments: Professionals versus Students," in *Handbook of Experimental Economic Methodology*, Guillaume R. Fréchette and Andrew Schotter (eds.), Oxford University Press.

Frederick, Shane, George Loewenstein, and Ted O'Donoghue (2002). "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature* XL (June):351-401

Glennerster, Rachel, and Kudzai Takavarasha (2013). *Running Randomized Evaluations: A Practical Guide*. Princeton, NJ: Princeton University Press.

Guala, Francesco (1999). "The Problem of External Validity (or 'parallelism') in Experimental Economics," *Social Science Information* 38(4): 555-73.

Guala, Francesco (2001). "Building Economic Machines: The FCC Auctions," *Studies in History and Philosophy of Science* Part A 32(3): 453-77.

Harrison, Glenn W. (1989). "Theory and Misbehavior of First-Price Auctions," *The American Economic Review* 79(4): 749-62.

Harrison, Glenn W., Morten I. Lau, and Melonie B. Williams (2002). "Estimating Individual Discount Rates in Denmark: A Field Experiment." *The American Economic Review* 92(5): 1606-17.

Harrison, Glenn W., and John A. List (2004). "Field Experiments," *Journal of Economic Literature* 42(4): 1009–55.

Hellerstein, Daniel, Nathaniel Higgins, and Michael Roberts (2015). *Options for Improving Conservation Programs: Evidence from Economic Experiments on the Design of Enrollment Mechanisms*. ERR-181. U.S. Department of Agriculture, Economic Research Service.

Holt, Charles A. (2006). *Markets, Games, and Strategic Behavior: Recipes for Interactive Learning*. Pearson Addison Wesley.

Horowitz, John K., and Kenneth E. McConnell (2002). "A Review of WTA/WTP Studies." J*ournal of Environmental Economics and Management* 44(3): 426-47.

Hughes, R. E. (1975). "James Lind and the cure of scurvy: an experimental approach," *Medical History* 19(4): 342–51.

Just, David R., and Joseph Price (2013a). "Default options, incentives and food choices: evidence from elementary-school children," *Public Health Nutrition* 16(12): 2281–88.

Just, David R., and Joseph Price (2013b). "Using incentives to encourage healthy eating in children," *Journal of Human Resources* 48(4): 855-872.

Lence, Sergio H. (2000). "Using Consumption and Asset Return Data to Estimate Farmers' Time Preferences and Risk Attitudes," *American Journal of Agricultural Economics* 82(4): 934-47.

Levitt, Steven D., and John A. List (2007). "What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?" *The Journal of Economic Perspectives* 21: 153-74.

Levitt, Steven D., and John A. List (2011). "Was there really a Hawthorne effect at the Hawthorne plant? An analysis of the original illumination experiments," *American Economic Journal: Applied Economics* 3 (1): 224–38.

List, John A. (2004). "The Nature and Extent of Discrimination in the Marketplace: Evidence from the Field," *The Quarterly Journal of Economics* 119(1): 49-89.

List, John A. (2008). "Informed Consent in Social Science," *Science* 322*:* 672.

List, John A., and Imran Rasul (2010). "Field Experiments in Labor Economics." NBER Working Paper No. 16062.

List, John A. (2011). "Why Economists Should Conduct Field Experiments and 14 Tips for Pulling One Off," *Journal of Economic Perspectives* 25(3): 3-16.

Messer, Kent D., Josh Duke, and Lori Lynch. (2014). "Applying Experimental Economics to Land Economics: Public Information and Auction Efficiency in Land Preservation Markets," *Oxford Handbook of Land Economics*. J. Duke and J. Wu, eds. Oxford University Press.

Messer, Kent D., and Alison Borchers (2015). "Choice for Goods Under Threat of Destruction," *Economic Letters* 135: 137-140.

Myerson, Roger B. (1981). "Optimal Auction Design," *Mathematics of Operations Research* 6(1): 58-73.

Palmieri, Paolo. (2009). "A Phenomenology of Galileo's Experiments with Pendulums," *The British Journal for the History of Science* 42(4): 479–513.

Payne, Colin R., Mihai Niculescu, David R. Just, and Michael P. Kelly (2014). "Shopper marketing nutrition interventions," *Physiology & Behavior* 136: 111-20..

Plott, Charles R. (1997). "Laboratory Experimental Testbeds: Application to the PCS Auction," *Journal of Economics & Management Strategy* 6(3): 605-638.

Price, Joseph, and David R. Just. (2015). "Lunch, recess and nutrition: Responding to time incentives in the cafeteria," *Preventive Medicine* 71: 27–30.

Roe, Brian E., and David R. Just. "Internal and External Validity in Economics Research: Tradeoffs between Experiments, Field Experiments, Natural Experiments, and Field Data," *American Journal of Agricultural Economics* 91(5): 1266-71.

Roth, Alvin E. (n.d.). "Laboratory Experimentation in Economics," *Economics and Philosophy* 2(2): 245-73.

Simon, Herbert A., and William R. Divine. (1941). "Controlling Human Factors in an Administrative Experiment," *Public Administration Review* 1(5): 485.

Smith, Vernon L. (1976). "Experimental Economics: Induced Value Theory," *American Economic Review* 66(2): 274-79.

Smith, Vernon L. (1982). "Microeconomic Systems as an Experimental Science," *American Economic Review* 72(5): 923-55.

Smith, Vernon L., Ronald R. King, Arlington W. Williams, and Mark Van Beoning (1993). "The Robustness of Bubbles and Crashes in Experimental Stock Markets," in *Nonlinear Dynamics and Evolutionary Economics*, Richard H. Day and Ping Chen (eds.). Oxford University Press.

Smith, Vernon L., Gerry L. Suchanek, and Arlington W. Williams (1988). "Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets," *Econometrica* 56(5): 1119-51.

Thaler, Richard H., and Cass R. Sunstein. 2009. *Nudges: Improving Decisions about Health, Wealth, and Happiness*. Penguin Books.

Thurstone, L.L. (1931). "The Indifference Function," *Journal of Social Psychology*, 2: 139-67.

U.S. Department of Agriculture, Farm Service Agency (2012). "Conservation Reserve Program Sign-Up 43 Environmental Benefits Index (EBI)."

U.S. Department of Agriculture, Farm Service Agency (2016a). "Conservation Reserve Program."

U.S. Department of Agriculture, Farm Service Agency (2016b). *The Conservation Reserve Program: 49th Signup Results.*

U.S. Department of Health and Human Services, Administration for Children and Families (2005). *Head Start Impact Study: First Year Findings*. Washington, DC.

U.S. Department of Health and Human Services, Administration for Children and Families (2010). *Head Start Impact Study. Final Report*. Washington, DC.

U.S. Department of Housing and Urban Development (HUD). *Housing Discrimination Against Racial and Ethnic Minorities.*

U.S. Securities and Exchange Commission, Office of Economic Analysis (2007). "Economic Analysis of the Short Sale Price Restrictions Under the Regulation SHO Pilot." Washington, DC.

Wallander, Steven, Paul Ferraro, and Nathaniel Higgins. (2017). "Addressing Participant Inattention in Federal Programs: A Field Experiment with the Conservation Reserve Program," *American Journal of Agricultural Economics.* April.

Wallis, W. A., and Milton Friedman. (1942). "The Empirical Derivation of Indifference Functions," in *Studies in Mathematical Economics and Econometrics in Memory of Henry Schultz*, O. Lange, F. McIntyre, and T. O. Yntema (eds.). Chicago, IL: University of Chicago Press.

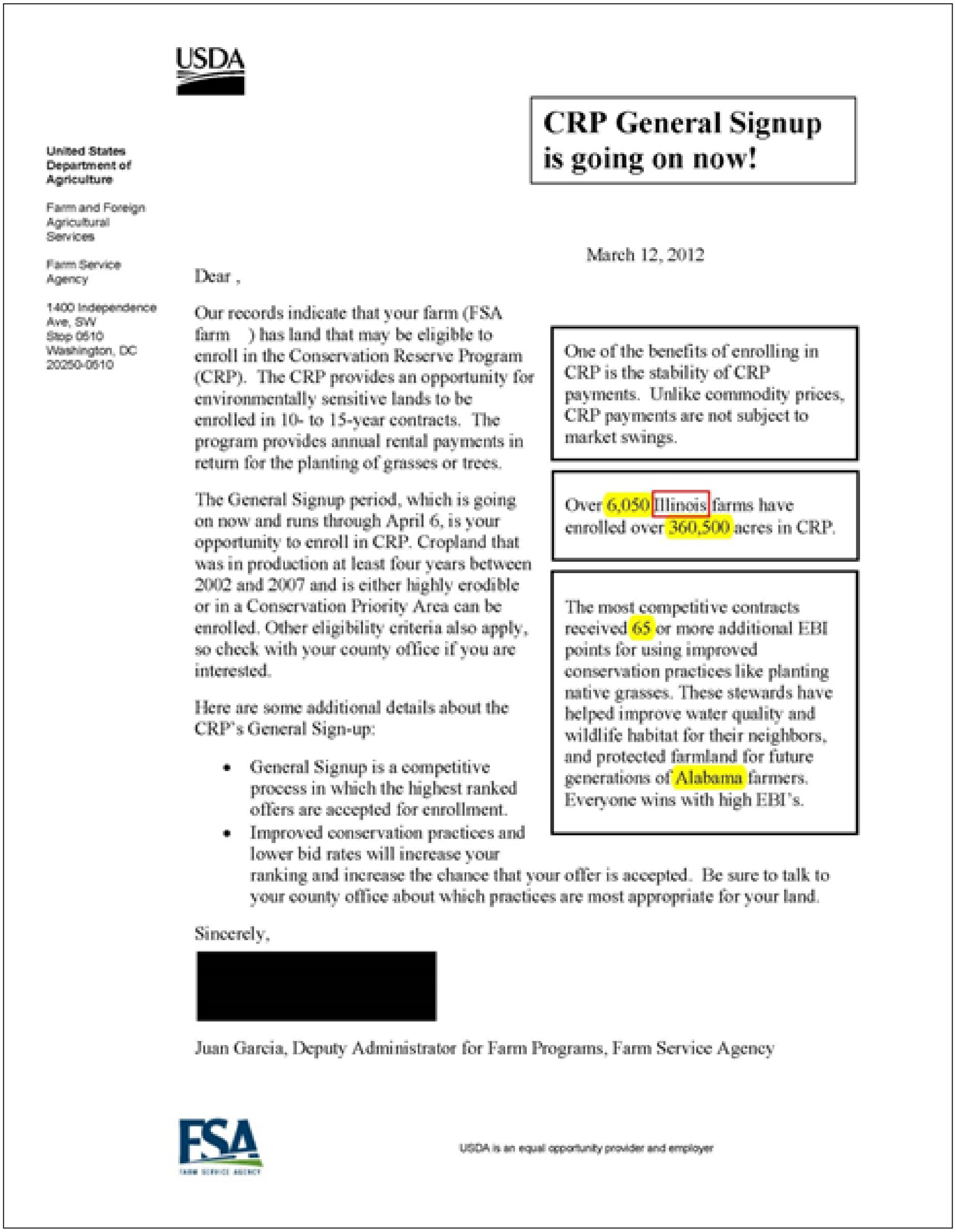
Warner, John T., and Saul Pleeter. (2001).”The Personal Discount Rate: Evidence from Military Downsizing Programs,” *American Economic Review* 91(1): 33-53.

Zarghamee, Homa., Kent D. Messer, Jacob R. Fooks, Willian D. Schulze, Shang Wu, and Jubo Yan. (2017.) “Nudging Charitable Giving: Three Field Experiments,” *Journal of Behavioral and Experimental Economics* 66. Feb.

Glossary

Asymmetric auction	An auction in which participants are not observationally the same. Systematic differences, such as in their reservation values, can be inferred by examining observable characteristics of the participants.
Bounded rationality	Classical economics assumes that individuals make "rational" decisions—they use all available information to compute and compare costs and benefits. However, this may not be how people really make decisions. Their information is limited, and they may not have time or cognitive capacity to fully use this available information.
Context	Context refers to the situation in which the experiment is conducted, and how similar that is to the situation in which the lessons of the experiment may be applied. For example, lab experiments often use generic language unrelated to any environment where participants make decisions; whereas field experiments center on an environment where the circumstances mimic actual decision made by the actual participants
Control	Control in this case refers to the degree to which factors can be held fixed during an experiment.
Ecological validity	When an experiment takes place in an environment that approximates natural conditions, it is said to have high ecological validity.
External validity	The ability to generalize to other contexts what is learned in a specific experiment.
Field experiment	An experiment that takes place in a natural setting, using participants similar to those to whom the lessons learned will be applied. There are several varieties of field experiments, including "artefactual," "framed," and "natural."
Homegrown values	A participant's own value—a value that existed before the experiment—for the target of an experiment (for example, for a physical object in an auction). In contrast, "induced values" are assigned to participants by the experimenter (for example, a dollar value assigned to "ticket" in a lab experiment).
Internal validity	The extent to which the estimation method identifies causal effect—if A happens, then B happens.
Laboratory experiment	An experiment conducted in an economics laboratory, which is usually a classroom environment with networked computer terminals and workstations for each participant in an experiment.
Policy experiment	While all experiments can be relevant to policymakers, policy experiments are meant to deliver evidence that is immediately actionable.
Representativeness	Do participants in the experiment behave like the relevant population?
Reservation value	The absolute "bottom-dollar" amount at which an auction participant is willing to make a deal.
Symmetric auction	Each participant in an auction is observationally identical.
Target environment	The naturally occurring environment in which policies are made and implemented.

Appendix 1—Example Letter from Experiment in Section Titled, "An Administrative Experiment: Encouraging Participation in the CRP"



USDA

United States Department of Agriculture

Farm and Foreign Agricultural Services

Farm Service Agency

1400 Independence Ave, SW
Stop 0510
Washington, DC 20250-0510

CRP General Signup is going on now!

March 12, 2012

Dear ,

Our records indicate that your farm (FSA farm) has land that may be eligible to enroll in the Conservation Reserve Program (CRP). The CRP provides an opportunity for environmentally sensitive lands to be enrolled in 10- to 15-year contracts. The program provides annual rental payments in return for the planting of grasses or trees.

The General Signup period, which is going on now and runs through April 6, is your opportunity to enroll in CRP. Cropland that was in production at least four years between 2002 and 2007 and is either highly erodible or in a Conservation Priority Area can be enrolled. Other eligibility criteria also apply, so check with your county office if you are interested.

Here are some additional details about the CRP's General Sign-up:

- General Signup is a competitive process in which the highest ranked offers are accepted for enrollment.
- Improved conservation practices and lower bid rates will increase your ranking and increase the chance that your offer is accepted. Be sure to talk to your county office about which practices are most appropriate for your land.

One of the benefits of enrolling in CRP is the stability of CRP payments. Unlike commodity prices, CRP payments are not subject to market swings.

Over 6,050 Illinois farms have enrolled over 360,500 acres in CRP.

The most competitive contracts received 65 or more additional EBI points for using improved conservation practices like planting native grasses. These stewards have helped improve water quality and wildlife habitat for their neighbors, and protected farmland for future generations of Alabama farmers. Everyone wins with high EBI's.

Sincerely,

[redacted]

Juan Garcia, Deputy Administrator for Farm Programs, Farm Service Agency

FSA FARM SERVICE AGENCY

USDA is an equal opportunity provider and employer

Appendix 2—Technical Details of the Experimental Estimation in Section Titled An *Administrative* Experiment: Encouraging Voter Participation in FSA County Committee Elections

Sampling Design

We randomized within two strata: (1) the election and (2) the type of household. The election is a locality variable—i.e. a variable that accounts for the local administrative area (LAA) in which the election is taking place. We refer to the type of household as either a "business" or an "individual." The identification of voters by these categories was based on available administrative data from FSA's Service Center Information Management System (SCIMS). A "business type" is coded in the system, separating out known business types—entities that have known legal structures—from individuals and those entities with unknown legal structures. We used these data to classify all voting entities that were known *not* to be individuals, such as partnerships, limited liability companies (LLCs), etc., into a category separate from the rest (the rest being "individuals," the presumed status of most entities in SCIMS). We also separated out voters who voted only in a single election from voters who voted in multiple elections. Because voters can operate farms in multiple LAAs, it is possible that individuals can vote in more than one election. This is not rare, but it is also not common. The analysis in this section deals only with the single-election voters, for ease of exposition.

Estimation

To estimate the effect of the treatments on voting behavior we use a multilevel model. The four treatments were formed by a factorial design, such that treatments one and two included postcard outreach, while treatments one and three included candidate names printed on the outside of each ballot. We estimate the effect of each facet of treatment—"postcards" (p) or "names" (n)—on voting behavior of household i (y_i) using a regression

$$y_i = \beta_0 + \beta_1 p_i + \beta_2 n_i + u_{iet}$$

where the coefficient β_1 is the effect of receiving postcards, β_2 is the effect of receiving a ballot with names printed on the outside, and u is an unobservable effect that varies over election (e) and entity type (t).[43]

The results of the regression are displayed below (and reported in summarized form in the main text).

	Estimate	Standard error	t value
β_0	0.105378	0.003115	33.82
β_1	0.024014	0.000518	46.36
β_2	0.004704	0.000518	9.08

[43]Several other regressions were run, including a four-dummy version that controlled for all treatments. In addition, we estimated a random effects model, estimated across the LAAs. These variants all yielded very similar results. Note that since the large number of observations yield vary small standard errors, other models designed to improve significance tests were not tested. In the interest of brevity, we just report the two-dummy model.