Purdue University Economics Department Working Paper No **1314**

Subject Pools and Deception in Agricultural and Resource Economics Experiments

Timothy N. Cason and Steven Y. Wu

Purdue University

August 2018

Abstract

The use of student subjects and deception in experiments are two controversial issues that often raise concerns among editors and reviewers, which might prevent quality research from being published in agricultural and resource economics (ARE) journals. We provide a self-contained methodological discussion of these issues. We argue that field professionals are the most appropriate subjects for questions related to policy or measurement, and students are the most appropriate subjects for scientific research questions closely tied to economic theory. Active deception, where subjects are provided with explicitly misleading information, has been avoided in the mainstream economics discipline because it can lead to a loss of experimental control, lead to subject selection bias, and impose negative externalities on other researchers. Disciplinary ARE journals may want to abide by these norms against deception to maintain credibility. Interdisciplinary ARE journals may have more flexibility, although it is important to provide guidelines to avoid too much reviewer-specific variation in standards. For ARE researchers, we suggest employing a deception-free experimental design whenever possible because we know of no field in which deception is *encouraged.*

JEL Classification: C90, Q10, Q30, Q50

Acknowledgements. For helpful comments, we would like to thank (without implicating) participants at the CBEAR-MAAP and WCERE conferences, and Simanti Banerjee, Carola Grebitus, David Cooper, Guillaume Fréchette, Nick Hanley, Leah Palm-Forster, Marco Palma, Sharon Raszap, Stephanie Rosch, Matt Rousu, Christian Vossler, and two anonymous referees. Wu gratefully acknowledges financial support from USDA-NIFA HATCH project IND010580.

1. Introduction

Experiments have become an essential methodology in the toolkit of agricultural and resource economists. They permit direct measurement of preferences and characteristics not usually observable in naturally-occurring markets and decisions, and they allow for controlled and exogenous variation to more easily infer causality. The growing popularity of experiments within the agricultural and resource economics (ARE) community is reflected by the fact that pre- and post-conference workshops specific to economic experiments have been conducted in recent Agricultural and Applied Economics Association (AAEA) annual meetings in Boston (2016) and Chicago (2017). Moreover, an increasing number of experimental papers are being published in *Environmental and Resource Economics*. Some recent examples include papers by Cason and de Vries (2018), Safarzynska (2018), and Mitra and Moore (2018), among others.

While there is an increasing acceptance of experimental research in the ARE community, two broad issues continue to pose challenges for agricultural and resource experimental economists, which can prevent solid research from being published. First, there is often skepticism concerning the use of student subjects rather than field professionals (e.g. Levitt and List 2007; Fréchette, 2015; Higgins, et al. 2017). Non-experimentalists sometimes make uninformed judgments about the value of such studies, while specialists might question the external validity of the findings. Moreover, given the growing popularity of field experiments over the last decade, there is also a tendency for peer reviewers and granting agencies to prefer "representative samples" drawn from the field. A prominent concern is that results derived from students may not generalize to field situations that involve "real world" decision-makers.

Second, the use of deception, whereby researchers might mislead subjects in the course of carrying out an experiment, can raise red flags to reviewers and editors. While the wider experimental economics community has long established protocols and norms on deception related matters, agricultural and resource economists often work across disciplinary boundaries via interdisciplinary research projects and grants. This can create confusion about the appropriate norms that might involve deception.

While it might seem odd to address subject pool matters and deception in the same paper, subject pool related issues and deception are, in our view, currently two of the most controversial issues in the ARE community. These two issues inspire many debates "behind the scenes" through

referee reports, at journals, and at funding agencies.[1] In some cases, the debates are explicit as in the case of a lively session at the 2013 AAEA annual meetings in Washington, D.C., and articles by Rousu, et al. (2015) and Colson et al. (2015) in AAEA journals. Moreover, Lusk (2018) discusses deception in a recent *Food Policy* article. With regard to subject pools, under the Environmental and Natural Resource Economics Program Priority Area (page 59) in the 2018 Request for Applications for the Agriculture and Food Research Initiative Competitive Grants Program, one of the highlighted research topics is to "….Examine and verify whether students are sufficiently representative of actual agents." Thus, subject pool related matters and deception are critical to both publishing and obtaining grant funding.

A unique contribution of our paper is that we offer a self-contained methodological discussion of subject pool issues and deception from the perspective of the larger disciplinary experimental economics community. Our hope is that the explicit articulation of these two issues from a disciplinary perspective can provide a basis for further dialogue, particularly with respect to how one might approach research where the norms regarding subject pools and deception are not always clear. We hope to provide some clarity on when student subjects are appropriate in experimental research, and what constitutes deception from the perspective of disciplinary economists (as opposed to what constitutes deception in other fields such as psychology). We also discuss how agricultural and resource economists might approach the issue of deception in order to maintain credibility with disciplinary peers, while not subjecting interdisciplinary researchers to the stringent anti-deception norms of disciplinary economists.

Though we discuss some interdisciplinary issues in passing, we do not provide an in depth discussion of all of the protocols and norms that might affect those who are engaged in interdisciplinary research. Although experiments are important in interdisciplinary research, it is simply not practical or feasible to go into detail about the wide array of possible protocols and norms across different interacting disciplines in a short journal article. Moreover, given the diversity of research that exists within the ARE community, it may not be practical nor desirable to establish unilaterally a single policy or viewpoint on subject pool matters and deception at the

[1] In the authors' experience, it is not unusual to receive referee reports that are broadly critical of the use of students as subjects rather than more specific aspects of a paper. For example, a report received by one of the authors did not mention anything specific to the paper but included the statement: "It is a matter of taste, I suppose, but I find experiments like these [using college students] uninteresting (unconvincing). I don't think the play of these games, where little is at stake, people don't have long to think about their strategies, and where the rules of the game are much clearer than in real life, tell us much about the real world."

community level. However, it might make some sense for specific journals to establish some guidelines depending on whether the journals cater to a disciplinary or interdisciplinary audience.

2. Student Subjects vs. Field Professionals

This section discusses some advantages and disadvantages of using field professionals rather than student subjects. Our main point is that the appropriate subject pool depends critically on the nature of the research questions being asked. Field professionals are most appropriate for a somewhat narrow but important set of questions, such as for measuring preferences and behavior for a specific population (e.g., farm operators) or when evaluating a specific conservation policy. Student subjects are relatively homogeneous compared to field professionals, and researchers can exert greater control in university laboratories.[2] This makes student-based experiments particularly valuable for theory testing and qualitative treatment comparisons.

Historically, carefully controlled economics experiments emerged from classroom demonstrations (Chamberlin 1948; Smith, 1962), which instructors conducted to illustrate and test theories. These demonstrations showed that experimental models of economic theories could be operationalized using students as decision-makers, which later led researchers at hundreds of campus-based laboratories to rely primarily on student subject participants (Svorencik, 2015). While economics experiments differ in many respects from experiments in psychology, both fields share this extensive reliance on student participants. Many other disciplines employ student volunteers in experimental research, making the undergraduate student one of the most intensely studied animals in science (Henrich et al., 2010).

In terms of providing qualitative insights and treatment comparisons about general economic theories, laboratory experiments with relatively homogeneous student subjects have distinct advantages over field experiments. Correct theories should be correct on all domains, including the simple and special cases where experimental investigations typically begin. Laboratory experiments also allow for more refined tests of theory following such initial tests, because they allow researchers to introduce incrementally more complex interventions to increase

[2] For example, student subjects are better able to follow neutral and abstract instructions relative to field professionals, who appear to find non-neutral framing helpful since it allows them to draw on their experience (Cooper et al, 1999; Alatas et al, 2009). Non-neutral framing is more likely to activate experimenter demand effects, however, and can lead to reduced control (Zizzo, 2010).

the level of stress on a theory (Just and Wu, 2009). The "external validity" of laboratory experiments arises from theory, not from any one specific lab session.[3]

Laboratory experiments can enhance external validity because the experimenter can manipulate numerous variables and factors to put stress on the theory and determine how sensitive the predictions of the theory are to context. A theory that can survive researcher-induced shocks or confounds can yield more confidence-inspiring predictions that are more likely to generalize across different contexts. As Camerer (2015, p. 252) puts it, experiments "are designed to contribute evidence about the *general* way in which individual characteristics, incentives, rules, and endowments influence economic behavior. Experiments contribute especially diagnostic evidence by virtue of extraordinary control over independent variables" (emphasis original). Laboratory experiments with student subjects are also (relatively) inexpensive. They can also stress test theories and carefully examine how context and salience affect results in considerably more controlled conditions than are available in field experiments.[4]

Many experiments are intended to provide qualitative and not quantitative comparisons across treatment conditions (Kessler and Vesterlund, 2015). For these types of investigations the focus is on treatment comparisons and hypothesis testing, so an optimal experimental design should minimize nuisance variation across treatments. Statistical inference is easier when variance is lower. A relatively homogeneous subject pool of undergraduate students allows the experimenter to minimize variation across subject characteristics that is not relevant for the theory. This increases the precision of estimates and statistical power for hypothesis tests, compared to a diverse set of field professionals who may differ considerably across age, education levels, and experience.

To illustrate this point, consider experimental research on non-human animals. Researchers in biology, neuroscience and many related fields have long recognized the advantage of using a relatively homogeneous animal model to minimize nuisance variation across treatments. For example, in many research studies using rat models, labs use the Fischer 344 Albino Rat, which originated in a Columbia University breeding colony nearly 100 years ago. Experimenters draw

[3] External validity refers to the ability of a causal relation identified in the experiment to generalize over subjects and environments (Fréchette, 2015).

[4] Al-Ubaydli and List (2015) point out that natural field experiments provide researchers with more control over who actually participates in an experiment because selection into participation is irrelevant. Subjects all participate because they do not realize they are in an experiment. Selection into participation cannot affect treatment comparisons, however, for laboratory experiments that employ random assignment to treatments.

conclusions on this animal model by treating these homogeneous rats differently, seeking insight into general biological or behavioral processes that apply to other rat breeds and in many cases other animal species. These conclusions are easier to draw because the genetic similarity within this particular colony helps isolate biological and behavioral differences arising from the applied treatments.

Of course, the widespread use of student subjects raises the natural question of whether researchers miss important insights when they focus on this non-representative type of subject. Fortunately, as discussed below, in most of the limited number of cases where experiments have made direct comparisons, conclusions do not differ significantly between student and non-student samples (Fréchette, 2015). Moreover, most research questions and theories in economics are intended to be general, and economics models almost never specify experience or background requirements for the agents that would render student subjects inappropriate. Nevertheless, a recent trend in ARE has emphasized a need to study decision-makers with relevant field experience, such as farmers and other landowners. The use of non-student subjects comes at considerable cost, however (Fréchette, 2016). For a controlled economics experiment, the salient financial rewards need to dominate any subjective costs or values associated with participation, as Smith (1982) defined in the *Dominance* precept. The high opportunity costs for field professionals generally requires substantially greater payments than are needed to motivate university students participating in convenient laboratories.[5] Another cost of using non-students comes from reduced replicability and experimental control. It is often difficult to bring non-student subjects to university laboratories where the researcher can strictly control the physical environment and information flows to increase replicability and control.

In contrast to laboratory studies that use student subjects, experiments using field professionals are often conducted in field settings and are referred to as artefactual or framed field experiments depending on whether neutral instructional terminology is employed (Harrison and List, 2004). Such experiments are most appropriate when the goal is to measure a specific characteristic or preference for this particular population. However, employing subjects that work in a specific region and subfield of an industry does not necessarily result in conclusions that generalize to other regions or parts of the industry or for the overall population. Thus, such studies

[5] Besides their high financial costs, in his handbook chapter Fréchette (2016) also identifies subject availability, replicability and limits to control as the four main disadvantages of studying representative samples and professionals.

may be internally valid for the specific region or industry but may not be externally valid across other settings or industries.

Indeed, when researchers study preferences or behavior of a more "relevant" population of field professionals, they usually do not employ a representative sample. To illustrate this point, consider a hypothetical study of the demand for higher education. Such a study could employ "field consumers" such as Ph.D. students in agricultural economics. This would lead to results that are likely to be internally valid if one were studying demand for graduate education among agricultural economics students. But the findings would likely not be externally valid if the researcher cared about generalizing the results to other populations, or for studying the overall demand for education in America. Similarly, many field experiments in agricultural economics recruit farmer subjects through university extension networks to increase participation, but this results in a selected class of farmers. Time and financial constraints, including costs of high incentive payments, may prevent the recruitment of a more representative sample.

Additionally, even if researchers study subjects from the population of interest, there is no guarantee that conclusions continue to hold for the *same* population after an exogenous shock alters the economic environment. Policy shifts or technology change can alter beliefs, which might lead to different behavioral responses. One would have to either repeat the experiment after the shock or rely on a robust theory that has been stress tested across different environments to make inferences about whether the conclusions continue to hold. Field experiments can be conducted on a highly relevant population and avoid subject selection by compelling participation, but they may still draw conclusions with limited external validity due to non-random selection of study sites (Allcott, 2015).

Nevertheless, some research questions require that experimental measurements be specifically focused on a field population of interest. For example, Roe (2015) seeks to compare the risk attitudes of farm owners in the U.S. to nonfarm business owners and the general population. In this survey-based study, farmers are less risk tolerant on average compared to other business owners but more risk tolerant than the general population. Herberich and List (2012) do not find significant differences between farmers and student subjects in their tolerance for "background risk," which is defined as risk that is not well-characterized by objective probabilities. The student and farmer experimental designs of that study differed in several important ways,

however, and the sample size was limited. This makes it difficult to draw definitive conclusions from this comparison.

Field professionals are also particularly relevant to study in choice experiments that measure how attribute levels affect stated preferences. Sampling field decision-makers allows the research results to be directly applicable for specific policies under consideration. For example, Kuhfuss et al. (2016) assess the potential benefits of offering a collective (agglomeration) bonus for reduced herbicide applications made by winegrowers in the South of France. The authors' research motivation was to gauge the potential of this bonus for improving the participation rate among this population in a specific agri-environmental incentive scheme. Use of the affected winegrowers as respondents clearly increases the external validity of the conclusions and policy recommendations for this particular program. The conclusions cannot be extended directly to settings outside the specific field context where the study took place, however, particularly in the absence of a general theoretical model (Levitt and List, 2007; Falk and Heckman, 2009).

Economics experiments have documented that preference and behavioral variance is greater in more diverse human subject pools, and how preferences depend on observable subject characteristics. For example, Andersen et al. (2010) compare risk and time preferences from a convenience sample of students at the University of Copenhagen to a representative sample of Danes. They find similar preferences on average between the field and laboratory samples, but much richer preference heterogeneity for the sample from the field. Moreover, they relate this variation in the representative sample to the greater variation in subject characteristics (place or residence, education and income level) for field participants. This variation is important to understand when the research goal is to measure the preferences and behavior for a specific population. But it is precisely this nuisance variation that researchers should seek to minimize when making treatment comparisons that are motivated by a theoretical model.

Maart-Noelck and Musshoff (2013) provide another example of a different form of behavioral variation with field professionals. They compare three different methods of measuring risk attitudes using German farmers, students, and Kazakhstani farmers. While they find inconsistencies in responses across the subject pools, students tend to respond more consistently across all three methods for risk attitude elicitation. In contrast, the responses of the farmers, both from Germany and Kazakhstan, tend to be more inconsistent across elicitation methods. Additional

research documenting how behavioral variance differs across student subjects and field professionals would be valuable.

Whether researchers draw misleading or inconsistent conclusions when using laboratory or field subjects is ultimately an empirical question. Unfortunately, direct subject pool comparisons such as reported in Maart-Noelck and Musshoff (2013) are relatively rare, and more systematic investigations are needed. The few direct comparisons of student and professional subjects have generally revealed relatively minor differences. In a recent survey Fréchette (2015) summarizes 13 papers that compare students and professionals within a standard laboratory environment. In nine of the 13 studies the conclusions do not differ across subject pools, and in the remaining four only one experiment finds behavior by professionals that is closer to the theoretical prediction.[6] For the other three studies, students conform more closely to theory or inexplicable differences exist across subject pools. In Carpenter and Seki (2011), for example, treatment effects are similar for student subjects and for the Japanese shrimp fishermen who do not traditionally share their income and operating expenses. The subsample of fisherman who pool their income and expenses in their field operations behave differently in the experiment compared to the other fishermen and the students.

Even fewer studies directly compare students and farmers. An important exception is Suter and Vossler (2013), who study the performance of ambient tax mechanisms among dairy farmers. Particularly because of the policy relevance of this research program, it is important to check the robustness of conclusions drawn for this regulatory mechanism for the target population since such farmers "may bring prevailing opinions, professional bias, and confounding norms and conventions" (page 92) to the experiment. If such factors turn out to be important, this would require changes to the theory and mechanism design. Fortunately, the main results regarding regulatory performance are robust to the different subject pools, at least in the aggregate. The authors observe some individual-level deviations, however, which they can relate to some farmer survey responses such as their attitudes towards environmental conservation.

This broad consistency of conclusions across subject pools arises in other samples as well, including for comparisons between (young) students and (old) retirees in Charness and Villeval

[6] For this study, Palacios-Huerta and Volij (2006) compare students and professional soccer players in zero-sum games with only mixed strategy equilibria. Wooders (2010) re-analyzes the same data and reaches the opposite conclusion--that students conform more closely to the mixed-strategy equilibrium.

(2009).[7] In Fréchette's summary (2016, page 472), he concludes that "cases where treatment effects (or comparative statics) are different when considering the standard [university student] subjects as opposed to other subjects are extremely rare." Therefore, when the research objectives are focused on making treatment comparisons and testing theory, the cost, availability and control advantages of student subjects all favor their use in laboratory experiments, and the existing evidence indicates that qualitative conclusions are unlikely to be affected. Field experiments with professionals certainly have an important complementary role, particularly to ensure robustness when policy conclusions are important or if the researchers wish to make measurements (e.g., preferences) for a specific population. Because of their lower cost and greater control, as a final practical point we also note that lab experiments with student subjects are very useful to refine experimental designs before they are taken into the field to study choices and preferences of professionals.

To conclude this section, we summarize our main points as follows:

- Experiments employing student or professional subject pools represent distinct methodological approaches that serve distinct purposes.
- Student subjects may be preferred for research questions that are related to testable hypotheses that emerge from theory. Students are a relatively homogeneous group, providing the researcher with more control over nuisance factors such as background context, experiences or biases that might be specific to only a small sub-group of field professionals from a particular industry. The reduction in variance from using students increases statistical power.
- Students are typically easier and less expensive to access, which facilitates replication and multiple sessions to facilitate stress testing of economic theories. Replicability significantly enhances scientific credibility. Theories that survive multiple stress tests are also more likely to be externally valid.
- Field professionals are generally preferred where the research questions concern specific policies (e.g. program evaluation), or where researchers are interested in measuring specific characteristics (e.g. risk preferences) of a particular population. Experiments with

[7] In a very recent study comparing students to a representative sample of Danes in a carefully controlled lab experiment, however, Fosgaard (2018) finds that students' cooperation levels decay much more to the selfish level predicted by standard economic theory.

field professionals generally have internal validity for such situations, but will not necessarily be externally valid for other populations and time frames.

- If researchers and stakeholders are primarily interested in the internal validity of the results for a specific program and population, especially with a goal of informing policy debates, then the use of field professionals may be preferred.

3. Deception

At a general level, experimental studies in several disciplines have used deception because in some circumstances, deceiving subjects about various facets of an experiment seems necessary in order to achieve certain scientific objectives. However, in principle, deception is rarely, if ever, *encouraged,* regardless of the field of study. In the most recently revised "Common Rule" guiding federally funded human subjects research, a new IRB exemption category for "benign behavioral interventions" that covers most economics experiments does not apply if deception is involved (Nichols et al., 2017). Thus, if a researcher wants to minimize compliance risk, a good starting point in formulating a research design is to ask whether the scientific objectives can be achieved without incorporating deception.

Unfortunately, things in practice are not always clear-cut. There are different norms and expectations across disciplines about what deception actually means and when exceptions can be made for deception. For agricultural and resource economists, this disciplinary inconsistency presents a particular challenge because many researchers within the field are increasingly working across disciplinary boundaries and/or publish in journals that have strong interdisciplinary emphases. This can create a confusing situation because agricultural and resource economists can receive mixed messages from interdisciplinary colleagues and journals who will lean on the norms and expectations of their own fields. Hence, it is no surprise that deception has been a relatively controversial issue for agricultural and resource economists in recent years. However, because the ARE field largely utilizes the tools, methodology, and language of economics and is widely considered a sub-field of general economics, we believe that the disciplinary norms from economics provide a useful benchmark. This is not to say that economic norms are "correct" in a philosophically optimal sense, nor is it to imply that all agricultural and resource economists should inflexibly adopt these norms. It is only to say that well-established norms already exist in economics and these norms and expectations have been reasonably successful in aligning

expectations regarding deception in the general economics community. Moreover, adhering to these norms may enhance credibility of the economics research conducted by ARE researchers and published in disciplinary ARE journals. In this section, we hope to provide a self-contained discussion of these norms and expectations. In addition, we will also provide some thoughts about when these norms might be relaxed to accommodate the specific needs of agricultural and resource economists.

Many economists believe that employing deception can lead to a loss of control and this has led to a long-standing objection to the use of deception in the general experimental economics community (Cooper, 2014). In fact, it is not uncommon for research papers that employ deception to be desk rejected in disciplinary economics journals. Of course, for increased control, researchers often employ neutral terminology to obscure the objectives and applied context of an experiment, so one can argue that subjects will always be "deceived" at some level. Hence, there will inevitably be grey areas when attempting to define deception and researchers often debate about exactly what constitutes deception.

With respect to the mainstream experimental economics community, Cooper (2014, p. 111) states that deception "…is generally taken to encompass instructions or materials that actively mislead subjects by stating or strongly implying something that is not true." Cooper further (p. 111) points out that deception, at least within the economics community, is "…a sin of commission rather than omission…" so that providing incomplete information is not generally considered deception but direct statements that mislead subjects is considered deception. This description of deception is not meant to be a precise definition that covers all circumstances but it does represent the general rules-of-thumb used in the experimental economics community.

The point of running an experiment is to obtain control over the subjects' environment in order to test theories and hypotheses with minimal variation in uncontrolled factors. Deception is not helpful to promote this primary goal of carefully studying implications of a specific economic theory. Theoretical predictions typically derive from the objective functions of the decision-makers, and a clean test requires subjects to understand correctly how their decisions affect payoffs (Cooper, 2014). This is also why an important feature of most economic experiments is the provision of monetary incentives. If deception is used repeatedly, then subjects may no longer trust the instructions in determining payment and the experimenter loses credibility with the subjects. Researchers will no longer know what incentives are motivating subjects.

This concern is potentially magnified because subject mistrust also affects other researchers using the same subject pool. Deception therefore imposes a negative externality on others, particularly on other research conducted in the lab that employed deception. Thus, the experimental economics community tends to be very conservative with regard to the use of deception and a strong anti-deception norm has developed as a precautionary safeguard against widespread loss of control over the incentives intended by experimenters. Jamison et al. (2008) showed that subjects who have been deceived in economic experiments tend to behave differently in subsequent experiments than those who have not been deceived. Most of the differences in behavior are driven by a selection effect where deceived subjects have lower rates of return to future experiments. In addition to selection, some deceived subjects who return also appear to answer questions related to risk inconsistently. Deception therefore can lead to bias and increased variability, threatening both internal and external validity even for other research that does not employ deception.

The most common types of deception in economic experiments, and which are objectionable to many economists, include not paying subjects according to the rules specified, telling subjects that they are playing against another subject when they are actually playing against a confederate of the experimenter (e.g. a computer or another human "subject" planted by the experimenter), or using randomization devices that differ from those stated (see Rousu et al., 2015, for a review). These all involve deliberate falsehoods, so the deception is clear. Potentially misleading subjects by omitting certain information is less clearly deception. A personal rule of thumb of one of the authors of this paper is to not omit any information that subjects might (based on the experimenters' subjective judgment) find objectionable if they learned about it later. One example of an omission that many experimental economists consider acceptable is to inform subjects that they are rematched randomly with new groups of subjects in each decision round, but omit informing them that this rematching only occurs within subgroups.[8] Another example is employing a "surprise restart" in which subjects learn about a new set of decision periods or tasks

[8] For example, subjects might be rematched into new pairs each round to play a 2-person game, with no information about their pair provided in order to avoid reputation formation and minimize any repeated game incentives. Subjects might participate in groups of 24 in the lab, but the rematching only occurs within two separate groups of 12 subjects. This increases the independence across the two groups who never interact. The detail about the subgroups is simply omitted from the instructions.

that provide opportunities to earn additional money, when these additional tasks were not announced earlier in the experiment instructions (see Wilson, 2016, for a careful discussion).[9]

Of course, exceptions are sometimes made and researchers include directly misleading statements in their experimental protocols. Cooper (2014, p. 113) suggests four circumstances, if they all hold simultaneously, under which such deception might be allowed: "(1) The deception does not harm a subject beyond what is typical for an economics experiment without deception; (2) The study is prohibitively difficult to conduct without deception; (3) Subjects are adequately debriefed after the fact about the presence of deception; (4) The value of the study is sufficiently high to merit the potential costs associated with the use of deception."

Deception might be less of a concern from the standpoint of imposing externalities on other researchers if the subjects are from a pool that is unlikely to be involved in future economic experiments. This is more likely to be the case if the experiment is conducted in the field rather than in a research lab, and the subjects drawn from the field setting will not interact with the researchers in any follow-up surveys or other later interactions. For example, Kröll and Rustagi (2016) provide strong evidence that experimental measurements of honesty, based on a commonly-used task in which payments are determined by privately-observed die rolls, correlate with natural honest behavior observed in markets. To measure market honesty the researchers bought milk from milkmen in India and measured how much water they used to dilute it. They needed individual measurements of die-roll honesty to correlate with the individual measurements of milk quality, so they employed Bluetooth-enabled dice that allowed observation of the individual die rolls. The milkmen subjects were deceived into believing that they observed their roll outcomes privately. Although the authors do not mention any steps to debrief subjects, this study meets Cooper's other 3 criterion for possible acceptability of deception. If these research subjects are highly unlikely to encounter experimenters in the future, the risk of negative reputational externalities is small. Of course, this risk will likely vary between large urban centers and small villages.

[9] Subjects could consider aspects of a surprise restart objectionable, however, if earnings opportunities in the second part depended on behavior in the first part in ways that were not revealed. In such cases the omission could be considered deceptive. For example, suppose that in a social dilemma such as a trust game or a public goods provision game subjects were regrouped in part 2 so that the most cooperative players in part 1 all interacted in part 2 and therefore earned considerably more in part 2. Had subjects known about this matching procedure they might have changed their part 1 behavior. The omission about matching groups' size described in the previous footnote is much less likely to have an impact on behavior.

Lusk (2018) provides a nuanced discussion of deception in field experiments using non-traditional subject pools where there is only a small or non-existent chance that deception will impose externalities on other researchers. He argues that reviewers and editors should avoid blanket bans on deception in these situations and should instead carefully weigh the benefits versus costs of employing deception. We agree with this point if the field respondents are highly unlikely to ever be invited to participate in another experiment. But if non-student subjects participate in multiple studies, such as when randomized control trials (RCT) are conducted repeatedly in the same geographic region, deception could lead to negative externalities and reduced control. Another important consideration is that, in some situations, particularly in developing countries, deception can lead to violence against the experimenters. In other cases, it is possible that local research assistants hired to run experiments may find the deception objectionable and work against the experimenter.[10]

Agricultural and resource economists often engage in interdisciplinary research, which can lead to situations where the lines between disciplinary norms are not always clear. This can create confusion because agricultural and resource economists might receive mixed messages from different disciplinary norms about deception. Rousu et al. (2015) discuss several key issues and sources of tension regarding deception, both within the agricultural economics community and across disciplines, particularly highlighting the differences between economics versus psychology. One of their points is that agricultural economists need to agree on a tighter, less ambiguous definition of deception. The call for a less ambiguous definition of deception is not surprising given the unique challenges that agricultural and resource economists face. We would like to make two points that will hopefully stimulate additional discussion.

First, it might be quite challenging to establish unambiguous definitions of deception at the level of the profession that is meant to apply to "agricultural and resources economists" generally. This is because agricultural and resource economists are a large and diverse group, and some engage in mostly disciplinary economics research while others engage in more interdisciplinary research. It is difficult to imagine a one-size-fits-all policy at the professional association level that would meet the needs of both groups of researchers. Agricultural and resource economists who mostly engage in disciplinary research face the possibility that their research might lose credibility

[10] We thank an audience member at the 2018 World Congress of Environmental and Resource Economists for bringing these points to our attention.

or be devalued if the standards deviate too far from the norms of economics, whereas interdisciplinary researchers may be unfairly handicapped because economists tend to be much more anti-deception relative to some other fields. Thus, applying a one-size-fits-all rule on deception may trigger a situation in which the community's norms may not fit either the disciplinary norms of economics or the norms of other interacting disciplines. This could be particularly problematic for early career researchers.

In our view, leaving some flexibility and discretion is not a problem, with the exception that it might increase the burden on individual researchers to educate themselves on the norms that govern their research. For example, if an agricultural economist decides to publish in a journal in another field outside of economics, it would be fruitful to spend some time learning the norms of that field and adhering to them. This is preferable to engaging in a discussion only among fellow resource and agricultural economists about whether those norms are appropriate, and then unilaterally defining new norms. Any attempt to rigidly define what is an acceptable level of deception only within ARE journals is likely to ultimately lead to frustration because norms can be so different across disciplinary and interdisciplinary outlets. Thus, we suggest that the ARE professional associations proceed with caution when attempting to construct a tight, unambiguous policy on deception for the field.

Our second point is that, at the journal level, it might make sense to be less ambiguous with regard to what constitutes deception, depending on the nature of the journal. Some journals are explicitly multidisciplinary (e.g. *Food Policy*) and are open to a wide variety of approaches and methods across different disciplines. One can imagine a situation where, without specific guidance about deception, the standards for making exceptions for deception can vary from reviewer-to-reviewer. This can be particularly challenging if one reviewer is from economics while another is from psychology because the norms can vary substantially across these two fields. Thus, it might be fruitful for these types of journals to provide, at a minimum, some general guidance about when exceptions will be made for deception so that the standards do not arbitrarily depend on the reviewers who are selected.

On the other hand, for disciplinary journals that primarily adopt the tools, methods and language of economics (e.g. *American Journal of Agricultural Economics (AJAE)* and *Environmental and Resource Economics (ERE)*), there is little downside to maintaining the deception norms of economics and a potentially significant downside to not doing so. A high

barrier exists to publishing papers that do not incorporate standard economic models or methods in these journals irrespective of whether deception is used. The downside to not adopting economics norms against deception could be lost credibility for the journals. These disciplinary journals have developed favorable reputations both among disciplinary researchers in ARE and among general economists. If they were to unilaterally abandon economics norms, these disciplinary journals may suffer some lost credibility by both agricultural economists who engage in disciplinary research and by general economists. As such, it may be appropriate for the disciplinary journals to adopt a policy on deception that is aligned with respectable mainstream economics journals, particularly *Experimental Economics*. This is a journal of the Economic Science Association (ESA), the premier international professional association for experimental economists.[11]

Any such no-deception policy requires editors *and reviewers* to recognize that omitting to inform subjects about some aspects of the experiment does not automatically constitute deception. All experiments involve some amount of omission; for example, most experiments do not reveal the research purpose and many employ neutral terminology to increase control. As noted already, previously unannounced additional parts of an experimental session, particularly if they are unrelated to early parts already completed, usually should not be considered deception. Another example is omitting information about the upper endpoint of the random price distribution in valuation studies using the Becker-DeGroot-Marschak mechanism. Such endpoint information is not relevant theoretically and it is not necessary to reveal it to subjects in order to avoid deception (Cason and Plott, 2014; Butler and Vossler, 2018). Omissions that are more likely to be objectionable include instruction wording that deliberately induces subjects to adopt incorrect beliefs, or that could lead subjects to not trust the experimenter subsequently.

One can also debate whether it makes sense to define an "optimal" or philosophically "correct" policy on deception. One problem is that the standards for making exceptions for deception within a discipline are as much an issue of norms and precedence as they are an issue of ethics or philosophical optimality. In experimental economics, the anti-deception bias is largely a norm that has emerged as a precaution against introducing confounds in subject pools and a loss of experimenter control. Other disciplines, such as social psychology, have their own norms and

[11] "…we only consider studies that do not employ deception of participants…" (https://link.springer.com/journal/10683, accessed 24 August 2018).

are often more permissible with regard to using deception but less permissible with regard to the use of monetary incentives to induce behavior. For example, psychologists have argued that deception is useful for obscuring the objectives of an experiment to avoid participants bias (Tyler and Amodio, 2015), although economists have countered that the distracting purpose suggested by the deception can lead to other biases (Offerman, 2015). One can make selective arguments to defend either norm on moral, philosophical or practical grounds. But in our estimation, these norms emerged because of some combination of differences in benefits and costs of deception across disciplines, precedence, and perhaps some historic path dependency. The fact that the norms for economists and psychologists evolved differently is not all that surprising, as Ortmann and Hertwiz (2002) and Roth (2001) point out that there are differences in the way the two disciplines approach research and the potential differences in public costs of using deception. Nonetheless, even the American Psychological Association code of conduct strongly discourages deception.[12] To be on the safe side, researchers ought to avoid deception whenever possible because we are unaware of any discipline in which deception is encouraged. There is only variation across disciplines about the circumstances under which it is acceptable. Therefore, "portable" experimental designs without deception that result in data that are publishable across all disciplines might be the safest starting point.

To be clear, we are not suggesting that the disciplinary norms in economics (or any other discipline) are immutable or sacred. Norms can change over time in response to better evidence or different practical considerations. However, it is also up to the researchers within a field to supply that evidence and to provide sound arguments for change. If agricultural and resource economists want to debate whether the disciplinary norms of economics are appropriate for a disciplinary economics field journal, the place to have that debate is not just within the small circle of agricultural economists but at general economics meetings or in economics journals. In the area of deception, the most appropriate outlet might be in the pages of *Experimental Economics* or at the ESA meetings.

Agricultural economists may have an advantage for engaging in this methodological debate. This is because their research topics can span across a multitude of fields, norms, and

[12] The code states "8.07(a) Psychologists do not conduct a study involving deception unless they have determined that the use of deceptive techniques is justified by the study's significant prospective scientific, educational, or applied value and that effective nondeceptive alternative procedures are not feasible" (APA, 2010, page 11).

subject pools, giving experimenters more flexibility in designing studies to investigate the potential benefits and costs of deception. This type of research would be of interest not just to agricultural and resource economists, but also to general experimental economists. If the evidence is compelling that deception norms should change, publishing the evidence in general economics journals could potentially lead to a shift in norms in the general experimental economics community and journals within the subfields should follow suit.

On another positive note, many agricultural and resource economists already appear to hold attitudes about deception that, with a few exceptions, are aligned with the expectations of the general experimental economics community. Colson et al. (2015) report the results of a survey of a sample of agricultural and applied economists who conduct experiments. The results suggest that the top five most "severe" forms of deception would also likely be unacceptable to the wider experimental economics community. These include deception that (1) cause physical or physiological trauma; (2) withhold promised payments; (3) cause subjects to purchase a mislabeled product; (4) provide subjects with false information about their performance; and (5) provide subjects with false feedback about other subjects' performance. The two types of deception that agricultural and applied economics do not consider "severe" but would likely be objectionable to the larger experimental economics community are (a) providing false information about the purpose of the study, and (b) using confederates who appear to be subjects but are actually working for the researcher. In short, it appears that only minor adjustments but no wholesale changes are needed to align views on deception with those of the parent discipline.

To end this section, we summarize our main points as follows:

- Deception is not "encouraged" in any discipline though there are different norms across disciplines about when exceptions are made. This can be confusing to agricultural and resource economists working across disciplinary boundaries.
- The wider experimental economics community tends to describe deception as a situation where subjects are *actively* misled by the experimenter rather than a situation where subjects are only provided with incomplete information.
- Economists discourage deception because it can lead to a loss of experimental control and introduce confounds, and also lead to subject selection bias. Deception can also impose negative externalities on other researchers.

- Deception induced externalities on other researchers might be less of a concern if subjects are drawn from a pool that is unlikely to be involved in future economic experiments. This is more likely to be the case for field experiments rather than laboratory experiments.
- Given the heterogeneous nature of ARE research, it may not make sense for the profession to unilaterally adopt a rigidly defined "official" policy or definition on deception. A one-size-fits-all-policy is unlikely to serve both disciplinary researchers and interdisciplinary researchers.
- At the journal level, interdisciplinary journals may want to provide tighter guidelines on when deception is acceptable given that different norms from different disciplines may clash and create confusion.
- Disciplinary journals may want to error on the side of adopting the norms of general economics in order to protect their reputation and credibility within the parent discipline.
- A good starting point for all researchers is to begin with a deception-free experimental design. Deception should only be incorporated if scientific objectives cannot be met any other way and the reasons for incorporating deception are consistent with the norms of the relevant discipline or journal.

4. Conclusion

Just as experiments provide an important complement to traditional empirical methods based on naturally-occurring data, different types of experiments provide complementary evidence that is useful for advancing the state of knowledge (Falk and Heckman, 2009). Lab experiments using university student subjects are most appropriate for addressing scientific research questions closely related to economic theory. Field experiments, including those conducted on farmers and landowners, are particularly appropriate for answering research questions relating to policy or for measuring specific characteristics of a field population. For such studies focused on measurement and policy, it would be valuable to have more research documenting any systematic subject pool differences to help identify their source—such as the differences attributable to neutral rather than framed experiment instructions (see footnote 2). This would help provide a clearer understanding of where obtaining a more representative subject pool is worth the cost.

The comparison of student subjects versus field professionals may also be relevant for studies that attempt to address emerging concerns about replication in the social sciences. Maniadis et al. (2017) provide a framework for the replication of experimental economics work, which depends on the statistical power of an experiment. Students and field professionals are likely to yield different benefit-cost ratios for achieving adequate statistical power in replication studies. Thus, future research might focus on how important these differences are in assessing the credibility of various types of economic experiments.

Both field and lab experiments in economics have traditionally forbidden deception, except for some mild forms based on the omission of benign details of the experimental environment. Other disciplines, especially psychology, have more relaxed norms regarding deception, and do not share this strict prohibition. Even so, the American Psychological Association's code of conduct discourages deception, unless there is a compelling reason to use it. Thus, we recommend avoiding deception unless the researchers have clearly studied the norms of the journals or disciplines that they plan to target. Leading agricultural and natural resource economics disciplinary journals, if they wish to maintain prominence within economics, may want to adhere to the wider experimental economics norms against deception.

Fortunately, agricultural and resource economists may have some comparative advantages in contributing to the larger methodological debate on deception. Agricultural and resource economists tackle research projects that involve a variety of fields, norms, and subject pools, including some subjects who will never again be invited to participate in another experiment. This gives them greater flexibility in conducting studies that specifically examine the costs and benefits of deception. Additional research and thoughtful consideration is needed to understand the implications of deception, especially in field experiments (see Ortmann, 2018, for a discussion). This includes the ethical issues arising from experimenting on unwitting subjects in a natural field experiment.

References

Alatas, V., Cameron, L., Chaudhuri, A., Erkal, N. and Gangadharan, L. 2009, Subject pool effects in a corruption experiment: A comparison of Indonesian public servants and Indonesian students. *Experimental Economics* 12, 113-132.

Allcott, H., 2015. Site selection bias in program evaluation. *Quarterly Journal of Economics* 130, 1117-1165.

Al-Ubaydli, O., List, J.A., 2015. Do natural field experiment afford researchers more or less control than laboratory experiments? *American Economic Review* (Papers & Proceedings) 105, 462-266.

American Psychological Association, 2010. Ethical Principles of Psychologists and Code of Conduct, APA 0003-066X, effective June 1, 2003, amended effective June 1, 2010, at https://www.apa.org/ethics/code/principles.pdf (accessed 24 May 2017).

Andersen, S., Harrison, G.W., Lau, M.I., Ruström, E.E., 2010. Preference heterogeneity in experiments: Comparing the field and laboratory. *Journal of Economic Behavior & Organization* 73, 209-224.

Butler, J.M., Vossler, C.A., 2018. What is an unregulated and potentially misleading label worth? The case of 'natural'-labeled groceries. *Environmental and Resource Economics* 70, 545-564.

Camerer, C.F., 2015. The promise and success of lab-field generalizability in experimental economics: A critical reply to Levitt and List. In *Handbook of Experimental Economic Methodology*, G.R. Fréchette & A.Schotter (eds.). Oxford Univ. Press, Oxford, UK, 249-295.

Carpenter, J., Seki, E., 2011. Do social preferences increase productivity? Field experimental evidence from fisherman in Toyama Bay. *Economic Inquiry* 49, 612-630.

Cason, T.N., de Vries, F.P., 2018. Dynamic Efficiency in Experimental Emissions Trading Markets with Investment Uncertainty. *Environmental and Resource Economics* forthcoming, https://doi.org/10.1007/s10640-018-0247-7

Cason, T.N., Plott, C.R., 2014. Misconceptions and game form recognition: Challenges to theories of revealed preference and framing. *Journal of Political Economy* 122, 1235-1270.

Chamberlin, E.H., 1948. An experimental imperfect market. *Journal of Political Economy* 56, 95-108.

Charness, G., Villeval, M.-C., 2009. Cooperation and competition in intergenerational experiments in the field and in the laboratory. *American Economic Review* 99, 956-978.

Colson, G., Corrigan, J.R., Grebitus, C., Loureiro, M.L., Rousu, M.C., 2015. Which deceptive practice, if any, should be allowed in experimental economics research? Results from surveys

of applied experimental economists and students. *American Journal of Agricultural Economics* 98, 610-621.

Cooper, D.J., Kagel, J., Lo, W. and Liang Gu, Q. 1999. Gaming against managers in incentive systems: Experimental results with Chinese students and Chinese managers. *American Economic Review* 89, 781-804.

Cooper, D.J., 2014. A note on deception in economic experiments. *Journal of Wine Economics* 9, 111-114.

Falk, A., Heckman, J.J., 2009. Lab experiments are a major source of knowledge in the social sciences. *Science* 326, 535-538.

Fosgaard, T.R. 2018. Cooperation stability: A representative sample in the lab. IFRO Working Paper No. 2018/08, University of Copenhagen.

Fréchette, G.R., 2015. Laboratory experiments: Professionals versus students. In *Handbook of Experimental Economic Methodology*, G.R. Fréchette & A.Schotter (eds.). Oxford Univ. Press, Oxford, UK, 360-390.

Fréchette, G.R., 2016. Experimental economics across subject populations. In *Handbook of Experimental Economics* Volume 2, J.H. Kagel & A.E. Roth (eds.). Princeton Univ. Press, Princeton, NJ, 435-480.

Harrison, G.W., List, J.A., 2004. Field experiments. *Journal of Economic Literature* 42, 1009-1055.

Henrich, J., Heine, S.J., Norenzayan, A., 2010. The weirdest people in the world? *Behavioral and Brain Sciences* 33, 61-135.

Herberich, D.H., List, J.A., 2012. Digging into background risk: Experiments with farmers and students. *American Journal of Agricultural Economics* 94, 457-463.

Higgins, N., Hellerstein, D., Wallander, S., Lynch, L. 2017. Economic Experiments for Policy Analysis and Program Design: A Guide for Agricultural Decision makers. Economic Research Report 236, United States Department of Agriculture-Economic Research Service, August.

Jamison, J., Karlan, D., Schechter, L., 2008. To deceive or not to deceive: The effect of deception on behavior in future laboratory experiments. *Journal of Economic Behavior and Organization* 68, 477-488.

Just, D.R., Wu, S.Y., 2009. Experimental economics and the economics of contracts. *American Journal of Agricultural Economics* 91, 1382-1388.

Kessler, J.B., Vesterlund, L., 2015. The external validity of laboratory experiments: The misleading emphasis on quantitative effects. In *Handbook of Experimental Economic Methodology*, G.R. Fréchette & A.Schotter (eds.). Oxford Univ. Press, Oxford, UK, 390-406.

Kröll, M., Rustagi, D., 2016. Shades of dishonesty and cheating in informal milk markets in India. SAFE Working Paper No. 134, Goethe University.

Kuhfuss, L., Préget, R., Thoyer, S., Hanley, N., 2016. Nudging farmers to enrol land into agri-environmental schemes: The role of a collective bonus. *European Review of Agricultural Economics* 43, 609-636.

Levitt, S.D., List, J.A., 2007. What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives* 21, 153-74.

Lusk, J.L., 2018. The costs and benefits of deception in economic experiments. *Food Policy*, forthcoming.

Maart-Noelck, S.C., Musshoff, O., 2013. Measuring the risk attitude of decision-makers: are there differences between groups of methods and persons? *Australian Journal of Agricultural and Resource Economics* 58, 336-352.

Maniadis, Z., Tufano, F., List, J.A., 2017. To replicate or not replicate? Exploring reproducibility in economics through the lens of a model and pilot study. *Economic Journal* 127, F209-F235.

Mitra, A., Moore, M.R. 2018. Green Electricity Markets as Mechanisms of Public-Goods Provision: Theory and Experimental Evidence. *Environmental and Resource Economics* forthcoming, https://doi.org/10.1007/s10640-017-0136-5

Nichols, L., Brako, L., Rivera, S.M., Tahmassian, A., Jones, M.F., Pierce, H.H., Bierer, B.E., 2017. What do revised U.S. rules mean for human research? *Science*, 357, 650-651.

Offerman, T., 2015. Discussion of 'Psychology and economics: Areas of convergence and divergence. In *Handbook of Experimental Economic Methodology*, G.R. Fréchette & A.Schotter (eds.). Oxford Univ. Press, Oxford, UK, 200-204.

Ortmann, A., Hertwig, R., 2002. The costs of deception: Evidence from psychology. *Experimental Economics* 5, 111–131.

Ortmann, A., 2018. Deception. In *Handbook of Research Methods and Applications in Experimental Economics*, A. Schram & A. Ule (eds.), forthcoming.

Palacios-Huerta, I., Volij, O., 2008. Experientia Docet: Professionals play minimax in laboratory experiments. *Econometrica* 76, 71-115.

Roe, B.E., 2015. The risk attitudes of U.S. farmers. *Applied Economic Perspectives and Policy* 37, 553-574.

Roth, A.E., 2001. Form and function in experimental design. *Behavioral and Brain Sciences* 24, 427–428.

Rousu, M.C., Colson, G., Corrigan, J.R., Grebitus, C., Loureiro, M.L., 2015. Deception in experiments: Towards guidelines on use in applied economics research. *Applied Economic Perspectives and Policy* 37, 524–536.

Safarzynska, K., 2018. The Impact of Resource Uncertainty and Intergroup Conflict on Harvesting in the Common-Pool Resource Experiment. *Environmental and Resource Economics* forthcoming, https://doi.org/10.1007/s10640-017-0193-9

Smith, V.L., 1962. An experimental study of competitive market behavior. *Journal of Political Economy* 70, 111-137.

Smith, V.L., 1982. Microeconomic systems as an experimental science. *American Economic Review* 72, 923-955.

Suter, J.F., Vossler, C.A., 2013. Towards an understanding of the performance of ambient tax mechanisms in the field: Evidence from upstate New York dairy farmers. *American Journal of Agricultural Economics* 96, 92-107.

Svorencik, A., 2015. The experimental turn in economics: A history of experimental economics. *Ph.D. dissertation*, University of Utrecht.

Tyler, T.R., Amodio, D.M., 2015. Psychology and economics: Areas of convergence and difference. In *Handbook of Experimental Economic Methodology*, G.R. Fréchette & A.Schotter (eds.). Oxford Univ. Press, Oxford, UK, 181-196.

Wilson, B.J., 2016. The meaning of *Deceive* in experimental economic science. In *Oxford Handbook of Professional Economic Ethics*, G. DeMartino & D. McCloskey (eds.) Oxford Univ. Press, Oxford, UK.

Wooders, J., 2010. Does experience teach? Professionals and minimax play in the lab. *Econometrica* 78, 1143-1154.

Zizzo, D.J. 2010. Experimenter demand effects in economic experiments. *Experimental Economics* 13, 75-98.