

Overall Assessment of the External Reviewers

The reviewers were asked to assess the following:

Experimental Design

Does the research propose appropriate methods to test a meaningful hypothesis?

All three researchers responded positively to this question. The reviewers asked for clarification on a few aspects of the design and proposed a few minor adjustments to the design as described more below.

Evaluate the match between the proposed treatments and the research question?

All three reviewers felt that the treatments would succeed in testing the research questions. The reviewers' requests for clarifications on the treatment design have been addressed as described in more detail below.

Evaluate the external validity of the experiment.

Overall, the reviewers felt that the external validity of the experiment was very strong given the focus on a farmer population and the careful design of the experimental auction to feature the key features of the CRP General Signup. The reviewers generally identified potential concerns about external validity that were very similar to the concerns already identified by the research team in the EDAP and acknowledged that the study had taken appropriate measures to address those concerns to the greatest extent possible.

Comment on the key features of the design such as the assignment of treatment, the statistical power, participant recruitment, and other features that you view as critical.

The reviewers felt that the sample size seemed large relative to other studies on conservation auction design, but they noted that this was appropriate given the power analysis. Reviewer 1 supported the use of waves of recruitment as a means for meeting the target number of participants.

Analysis Plan

Are the statistical tests for treatment effects appropriate?

The reviewers responded positively to this question. Some of the reviewers suggested alternative or additional tests for consideration. The research team's response to those suggestions is detailed below.

Have the researchers left out any tests of analysis that should be conducted?

The reviewers suggested greater use of control variables in the estimation of the treatment effects, a suggestion which the research team feels is not worth the added response burden. Additional detail on this is provided below.

Changes and Additions Made in Response to the Reviewers

Clarifying the motivation for testing student and farmer populations

Reviewer 2 did not understand the need for testing the outcomes of students versus farmers. In part this is because this reviewer has a strong prior belief that findings from a student population would have external validity for the given research questions.

Response: Supporting Statement A was written to provide additional clarification on this point. In addition, the ERS white paper on the issue of students as subjects for experimental studies is included with the ICR to provide greater clarification on why this hypothesis is being tested.

Information on the ranking (EBI) calculation

Review 1 requested clarification on how the EBI is calculated and how the EBI cutoff is determined. This comment has important implications for program participants as well.

Response: The research team focused heavily on this comment and how it relates to external validity. The goal in developing the study interface is to provide participants with the same type of prior information on scoring and ranking that they would receive in the CRP General Signup or similar conservation auctions. (This study is explicitly not focused on the hypothesis investigated in one strand of the literature on conservation auctions, which looks at the impacts on auction outcomes of withholding the scoring and ranking criteria.) The data collection instrument was edited to explicitly include all additional EBI points for practices for participants as well as a graphical indicator of the relationship between the EBI and the probability of acceptance. This design reflects the sort of information that is available to program participants during a regular general signup.

Information on auction clearing rule:

Reviewer 1 requested clarification on how the auction clears. Reviewer 3 raised related concerns about the potential for portfolio effects if all three auction rounds are used for calculating a payment. Reviewers 2 and 3 both suggested clearing the auction by randomly selecting one of the rounds for purposes of making the payments and communicating this to participants. Reviewers 1 and 3 each raised a related concern, asking what participants know about the distribution of field characteristics for their competitors.

Response: The problem of clearing the auction in a way that maximizes external validity was one of the major design challenges for this study. The CRP General Signup is a closed-offer, single-round auction in which participants face considerable uncertainty about the final acceptance rate or “cut-off EBI.” That auction typically clears several weeks after the signup period closes. The overall design of this study in which participants do not observe any rounds of the auction clearing directly reflects this aspect of CRP General Signup. The challenge in this study, then, was to provide sufficient information to participants for formation of some level of understanding about how the auction would clear. The instructional materials were redesigned to provide clarity on the use of a single round for the payment. The software added a visual scale to give a heuristic device for individuals to understand the relative likelihood of acceptance without using probability terms. The research team felt that this most closely reflects the

type of uncertainty in an actual option in which participants do not know how many offers will be made but they may have a broad idea of their changes based on recent auction. The study also now includes random selection of one round for making payments and communicates this to participants in the instructional materials.

Scale of incentives relative to the actual CRP

Review 1 asked about the incentives in the study compared to the incentives for CRP. Reviewer 2 raises a similar issue while noting that in the actual CRP producers face a number of non-pecuniary incentives.

Response: Supporting Statement A was written to include a summary of this comparison when discussing payment levels. The research team also agreed with reviewer 2 that non-pecuniary incentives are important in CRP but cannot be captured in a lab experiment.

Literature review

Several of the reviewers requested additional literature citations.

Response: The ICR includes a literature review.

Numeracy or deliberative thinking task

Review 1 asked if the team planned to conduct a numeracy or deliberative thinking task.

Response: After considering a way to address this issue while minimizing response burden, the research team decided to add a few questions after the last round of the study. These questions are simple agree/disagree questions about comprehension and strategy in the study. In addition to providing some statistics to help evaluate comprehension, we think this will also be helpful for interpreting and differences between the farmer and student populations.

Probability of Acceptance

Reviewer 2 suggested providing information to participants on their expected probability of acceptance using a color ramp.

Response: The design of the survey instrument now includes such a visual guide.

Acreage Endowments

Reviewer 2 asks “why the odd choice to endow students with a single acres and farmers with five acres?”

Response: The supporting statements now include discussion (in the section on participant payments) that explain this decision in more detail. The reviewer is correct in noting that the difference in payment adds a factor that could contribute to different outcomes for the two populations. Rather than being a source of bias, as suggested by the review, that is an intentional choice related to reasons for comparing these two populations. A major motivation for conducting a great deal of experimental economics on student population is the lower cost of participation and incentive payments. In other words, providing exactly the same payment levels to students and farmers would actually bias the results toward the null hypothesis of no difference in outcomes by failing to replicate one of the standard differences between lab experiments with student and non-student populations. Based on that

assessment, the research design explicitly includes a meaningful difference in payment levels. To maintain the relative marginal tradeoffs within the auction design (i.e.: the change in relative ranking for a given practice choice or percent reduction in payment), the incentives are simply scaled up and down based on the number of acres.

Reviewer Comments and Suggestions that Required Only Clarification or that the Research Team Did Not Incorporate

Design Comments and Suggestions

Reason for 2x2 design:

Reviewer 1 asked for clarification about the use of a 2x2 treatment design.

Response: The 2x2 design without testing for an interaction effect maximizes the statistical power of the sample by splitting the sample in half for both treatments while keeping the assignment of those treatments independent from each other.

Increase the Number of Rounds

Reviewer X suggest tincreasing the number of rounds and discusses possible portfolio effects.

Response: random selection of round reduces likelihood of portfolio effect. Multiple rounds would increase respondent burden. While there would be a gain in power, that gain is diminishing within increasing rounds since it is only variation in the underlying field characteristics being added given that the treatment is constant across rounds for each participant.

Treatment suggestions

Additional Default Treatment

Reviewers 1 and 3 suggested that the study add a second default treatment at lower offer improvement level.

Response: In terms of adding an additional treatment, the research team does not feel that the study has sufficient power for additional treatments. In terms of selecting a mid-scoring default as opposed to the high scoring default, the research team is proceeding with the assumption that the highest scoring default would have a larger treatment effect, since it is the furthest point from the Nash Equilibrium offers and therefore is a more salient anchor. This means that the treatment effect on the high-scoring default is expected to be easier to detect, if there is such an effect, than the treatment effect on a mid-scoring default.

Analysis suggestions

Rank sum test

Reviewer 1 suggests the use of a rank sum test but also notes that the large sample size in this study probably justifies not using a rank sum test.

Response: While a rank-sum test would be an interesting non-parametric alternative to the proposed tests, it generally has lower statistical power than a parametric test such as a t-test. The rank-sum test is sometimes favored if a study has small samples sizes or skewed data. As the reviewer notes, this study will have a large sample size. The research team feels that it is appropriate to retain the current tests as the primary test and to only include a rank-sum test as a robustness check, especially if the response on rental rates or practices are more skewed than predicted by the Nash Equilibrium model.

Control variables

Reviewer 1 asked for clarification about any control variables that would be included.

Response: The main reason for including such variables would be to increase the precision of the treatment effect estimates. These variables would not reduce bias or improve consistency since the random assignment of treatment and the large population provides fairly strong assurance that the treatment will be uncorrelated with any control variables. In addition, there are two key factor that also make this difficult to implement. First, the literature on conservation auctions does not reveal any such variables that consistently predict offer structure with high precision outside of the auction design characteristics that are already used for inducing variation in incentives. In addition, even if such variables could be found, it would increase respondent burden to collect such information, which is difficult to justify if there is only a possible and perhaps small increase in efficiency.

Order effects using control test

Reviewer 2 suggests that if the team has sufficient statistical power, they could test for order effects.

Response: The research team feels that the study does not have the statistical power to implement this test. However, as with the rank sum test suggest by reviewer 1, the research team does feel that this would be an appropriate “exploratory” test. Since such a test would likely be underpower, any finding based on this test would simply be suggestive for future lines of research.

Heterogeneous effects

Reviewer 3 suggest that the team interact default treatment with EBI endowment to test for heterogenous treatment effects.

Response: This is an interesting suggestion since the anchoring effect of a fixed default is likely to vary with respect to the “distance” of the default from the Nash equilibrium offers. However, the research team felt that the study does not have the statistical power to conduct this test. The team felt it would be appropriate to conduct this as an exploratory analysis. And statistically significant finding on this test would be a justification for future research designed explicitly to test for such heterogenous effects.