THE EXPERIMENTAL SCIENCE OF ECONOMIC BEHAVIOR: TESTING THEORIES OF PARTICIPATION, VALUATION, AND INNOVATION

By

Stephen Nicholas Morgan

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Agricultural, Food and Resource Economics – Doctor of Philosophy

2018

ABSTRACT

THE EXPERIMENTAL SCIENCE OF ECONOMIC BEHAVIOR:
TESTING THEORIES OF PARTICIPATION, VALUATION, AND INNOVATION

By

Stephen Nicholas Morgan

Agricultural production takes place in the presence of a complex web of institutional and market settings that can create countervailing incentives for producers. Economists have long been concerned with understanding the causal mechanisms driving producer decision-making and behavior. Experimental methods provide one way to disentangle the effects of specific policies and interventions on economic and producer behavior. This dissertation leverages experimental methods to better understand producer behavior in areas of stakeholder participation, technology valuation, and agricultural innovation.

In essay 1, we design an experiment to identify the marginal impact of a comment mechanism on individual public goods provision. First, we introduce a novel comment mechanism where individuals can provide unstructured feedback to influence a third-party rule-maker whose sole task is to determine a minimum contribution rule (MCR) at the beginning of the game. Second, we implement a costly probabilistic sanctioning mechanism for individuals who do not contribute at least the MCR to the public good. Our results suggest that when the exogenous enforcement mechanism is present, the comment-based participation mechanism has a large positive effect on players' contributions to the public good and on their compliance with the MCR. The comment-based participation mechanism has a large positive effect on the MCR set by the rule-maker; players, in turn, respond to an increase in the MCR by making larger contributions to the public good. Overall, comment-based participation mechanisms have similar effects on contribution levels to those of nonbinding voting mechanisms.

Essay 2 contributes to the broader literature on extension services and farmer valuation of new technologies. We use a Becker-DeGroot-Marschak (BDM) mechanism to elicit farmer WTP for new bean seed technologies under two different lead-farmer extension treatments in Tanzania. The first treatment uses the planting of demonstration plots to showcase the new technologies within a village. The second treatment includes the same demonstration plot approach but adds the distribution of trial packs of the improved inputs that allow some farmers to test the technologies on their own plots of land. Despite successfully raining awareness of new agricultural technologies, our results suggest that neither extension treatment significantly affects WTP for the improved bean technologies we consider. Additionally, we find that farmers are willing to pay more for bean seeds that are pre-treated than for those they have to treat at home suggesting that lead-farmer extension might deliver more value through the provision of basic services than in demonstration roles.

Essay 3 tests the effects of increased rates of technological innovation on technology adoption behavior. Using a lab experiment, we implement high and low innovation rate treatments to test for behavioral effects on the probability of adopting a new technology and the number of technologies adopted. Our decision environment incorporates technologies that generate stochastic returns, partial irreversibility through fixed arrival costs, and uncertain arrivals and returns to future innovations. We find mixed results for the effect of increased rates of innovation as a driver of delay. Participants in the high innovation rate treatment group are more likely to adopt a new technology as soon as it arrives and are less likely to continue using a current technology within a given round. However, when comparing across a common set of technologies, subjects in the high innovation treatment adopt fewer innovations overall.

For my parents, Ronnie and Jennifer Morgan, and my grandmother, Janis Sironko.
With all my love.

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Nicole Mason for being an incredible role model, mentor, and friend. Thank you for freely sharing your time, talents, and advice over the past four years to help me develop my ideas and skills as a researcher, teacher, and writer. I would like to thank Dr. Robert Shupp for your willingness to provide feedback on all of my ideas and for teaching me the fundamentals of designing and running economic experiments. I also thank my committee members Dr. Robert Myers and Dr. Rachel Croson for their support and suggestions that improved my dissertation research at every stage.

I would also like to thank the entire faculty and staff of the Department of Agricultural, Food, and Resource Economics (AFRE) for providing me with the skills and support needed to conduct my research and engage with our field. Special thanks to Dr. Roy Black for many helpful discussions about research design and farmer behavior. I also thank Nancy Creed, Ashleigh Booth, and Debbie Conway for all of their work to make sure I never ran out of time and money.

Thank you to Dr. Evan Drummond for inviting me into the field that I love and for your willingness to provide perspective on all of my hard decisions. I aspire to be the kind of mentor and friend to my students that you have been to me. Great professors not only teach students but inspire them. Dr. Leann Brown, Dr. Aida Hozic, and Dr. Richard Scher have never failed to inspire me with their intellect, kindness, and compassion inside and outside of the classroom.

Mutual hardship often helps you find the best of friends. Thanks to Sam Padilla for being hilarious, endearing, and always being ready to celebrate our successes and my failures. Thanks to Sophia Tanner for teaching me almost all of the econometrics that I know and for all of the adventures across campus. Thanks to Mary Doidge for allowing me to work in your office, for never putting up with my nonsense, and for introducing me to Canadian culture, eh? Thanks to

Asa "Snacks" Watten for all of the weird snacks. Graduate school wouldn't have been nearly as fun without you all.

Thank you to Dr. Christopher Marth, Ph.D., J.D. for being a brilliant friend and always providing a source of colorful quality commentary on my life. Your journey through vocational school has truly been a source of inspiration.

Finally, none of this would have been possible without my parents, Ronnie and Jennifer Morgan. Thank you for your unwavering love, support, and encouragement. Not only were you willing to learn with me about all of the ins and outs of graduate school, but you taught me the value of hard work, laughter, and persistence each and every day. I couldn't have done any of this without you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

APE	Average partial effect
ARI	Agricultural research institute
BDM	Becker-De Groote-Marschak
C	Comment only
CE	Comment with enforcement
CIAT	International Center for Tropical Agriculture
DP	Demonstration plot
DPTP	Demonstration plot and trial packs
E	Enforcement only
EPA	Environmental Protection Agency
FFS	Farmer Field Schools
FIPS	Farm Input Promotions
HYV	High-yielding variety
KG	Kilogram
LHS	Left-hand side
MCPR	Marginal per capita return
MCR	Minimum contribution rule
MSU	Michigan State University
NGO	Non-governmental organization
NJ	Njano Uyole
OLS	Ordinary least squares
R&D	Research and development

RCT	Randomized controlled trial
RHS	Right-hand side
T&V	Training and visit
TFP	Total factor productivity
TSH	Tanzanian Shilling
U96	Uyole 96
VBAA	Village-based agricultural advisor
VCM	Voluntary contribution mechanism
WTP	Willingness-to-pay

INTRODUCTION

Experimental methods have become a workhorse in the agricultural economics discipline to better understand the observed behavior of consumers and producers. Early work focused primarily on issues of consumer demand, particularly on the elicitation of willingness-to-pay (WTP) and willingness-to-accept (WTA) for different food products and food product attributes (e.g. Shogren et al. 1994; Hayes et al. 1995; Roosen et al. 1998; Lusk et al. 2001; Lusk, Norwood and Pruitt 2006). Experimental research on supply-side issues emerged in agricultural economics through a need to better estimate the role of farmer risk aversion in understanding on-farm decision-making in a developing country context (e.g. Binswanger 1981; Schechter 2007). Experimental studies of market and institutional designs to serve producers have since proliferated to include, for example, the design of auction mechanisms for natural resources and conservation (e.g. Cummings, Holt and Laury 2004; Palm-Forster et al. 2016), the role of cognitive ability in technology adoption (Barham et al. 2018), and farmer preferences for improved inputs (Waldman, Kerr and Isaacs 2014).

This dissertation consists of three essays using a mixture of lab and field experiments to analyze issues confronting agricultural producers in areas of stakeholder participation, technology valuation, and technology adoption. The first and third essays use decontextualized lab experiments to test behavioral hypotheses in a controlled environment. This is important as I focus on the analysis of previously untested comment and innovation mechanisms, respectively. The second essay implements real auctions with smallholder farmers in Tanzania. Results from this dissertation will help policymakers understand possible impacts of different institutional and market arrangements on producer behavior.

The first essay (Chapter 2) is titled "Stakeholder Comments, Contributions, and Compliance: Evidence from a Public Goods Experiment". In this essay, I use a public goods experiment with a voluntary contribution mechanism to estimate the effect that eliciting stakeholder participation through open-ended comments might have on individual contributions to public goods and compliance with final rules and policies. Stakeholder engagement via comments are popular forms of deliberative democracy in the agricultural and natural resource sector (Gardner 2009; Yackee 2006; Crow, Albright and Koebele 2016), but previous public goods experiments have largely confined the analysis of participatory mechanisms to majority-rule voting (Kroll, Cherry and Shogren 2007; Messer et al. 2007). The comment treatment is introduced to the public goods game both with and without an exogenous enforcement mechanism to better understand any interaction effects that may exist between participatory and policing institutions. Outcome measures include total individual contributions to the public good and two measures of noncompliance Results suggest that comments have a positive effect on individual contributions and compliance, conditional on the presence of a complementary enforcement mechanism.

In addition to contribution and compliance behavior, I also allow for a third-party to receive stakeholder comments before determining a contribution rule to the public good. In this environment, subjects share the same payoffs (and costs) to public goods, and all individuals making contribution decisions could be made better off with a higher contribution rule that everyone adheres to. I find evidence that policymakers respond strongly to the comments they receive in this experiment by raising recommended contribution rules closer to the Pareto optimal level. This effect is independent of the existence of any enforcement mechanism. Results from this essay should be of considerable interest to regulators and policymakers who are often required to extend comment opportunities to diverse interest groups.

The second essay (Chapter 3) is titled "Do Different Extension Approaches Affect Smallholder Farmers' Willingness-to-Pay for New Agricultural Technologies? Experimental Auction Results from Tanzania". Encouraging the widespread adoption and use of new on-farm technologies is an important part of productivity-led strategies to promote agricultural transformation. While many interventions have been designed to promote adoption through extension and education, little is known about how these efforts influence farmer willingness-to-pay (WTP) for new technologies in the marketplace. In this essay, I use a Becker-DeGroot-Marschak (BDM) mechanism to elicit farmer WTP for new bean seed technologies under two different extension treatments (Melkani and Mason 2018) in Tanzania. The first treatment uses the planting of demonstration plots to showcase the new technologies within a village. The second treatment includes the same demonstration plot approach but adds the distribution of trial packs of the improved inputs that allow some farmers to test the technologies on their own plots of land. I find little evidence to suggest that either treatment significantly affects WTP for the improved bean technologies considered.

An additional contribution of this paper is that I am also able to assess the effect of different bean technology attributes on producer WTP. Specifically, I focus on differences in WTP for bean seed including a chemical seed treatment - one version of which comes pre-applied and another version of which must be applied by the farmer. This difference is a good approximation to understand how smallholder producers might value the provision of basic agricultural services (e.g. seed treatment) in rural communities. Results suggest that farmers are willing to pay more for bean seed that is pre-treated than for seed that they have to treat at home. These results should be of interest to policymakers, extension agents, donor organizations, and the private firms seeking to market their new technologies in developing countries.

The third and final essay (Chapter 4) is titled "Agricultural Innovation and Technology Adoption: Strategic Delay Reconsidered". Despite vast gains during the 20th century, recent evidence suggests that the rate of growth of agricultural productivity is declining due to decreases in overall research and development (R&D) spending (Alston et al. 2010; Ball, Schimmelpfennig and Wang 2013; Pardey et al. 2015). Calls for corrective action (Cai, Golub and Hertel 2017) suggest that a boost in R&D funding would accelerate innovation processes and increase the number of productivity-enhancing technologies that will become available to farmers in the future. My research question focuses on how a change in the rate of innovation, measured by the speed of arrival of new technologies, affects individual adoption decisions over new technological innovations. I situate the problem in a discrete real-options framework where uncertainty over the arrival and the returns to a new innovation may generate significant incentives for an individual to delay adoption to a later period.

Data is collected using a lab experiment that implements high and low innovation rate treatments. The decision environment incorporates technologies that generate stochastic returns, partial irreversibility through fixed arrival costs, and uncertain arrivals and returns to future innovations. Additionally I control for risk preferences using a measure of each participant's relative risk aversion that was elicited through a lottery choice experiment (Holt and Laury 2002). Outcome measures of interest include the probability of adopting a new technology and the number of technologies adopted. Results suggest that a higher rate of technological innovation makes individuals more likely adopters of new technologies as soon as they become available, however they adopt fewer new technologies over time.

REFERENCES

REFERENCES

Alston, J.M., M.A. Andersen, J.S. James, and P.G. Pardey. 2010. *Persistence pays: U.S. agricultural productivity growth and the benefits from public R&D spending*. New York: Springer.

Ball, E., D. Schimmelpfennig, and S.L. Wang. 2013. "Is U.S. Agricultural Productivity Growth Slowing?" *Applied Economic Perspectives and Policy* 35(3):435–450.

Barham, B.L., J.-P. Chavas, D. Fitz, and L. Schechter. 2018. "Receptiveness to advice, cognitive ability, and technology adoption." *Journal of Economic Behavior & Organization* 149:239–268.

Binswanger, H.P. 1981. "Attitudes Toward Risk: Theoretical Implications of an Experiment in Rural India." *The Economic Journal* 91(364):867.

Cai, Y., A.A. Golub, and T.W. Hertel. 2017. "Agricultural research spending must increase in light of future uncertainties." *Food Policy* 70:71–83.

Crow, D.A., E.A. Albright, and E. Koebele. 2016. "Public Information and Regulatory Processes: What the Public Knows and Regulators Decide: Public Information and Regulatory Processes." *Review of Policy Research* 33(1):90–109.

Cummings, R.G., C.A. Holt, and S.K. Laury. 2004. "Using laboratory experiments for policymaking: An example from the Georgia irrigation reduction auction." *Journal of Policy Analysis and Management* 23(2):341–363.

Gardner, B.L. 2009. *American Agriculture in the Twentieth Century: How It Flourished and What It Cost.* Cambridge: Harvard University Press.

Hayes, D.J., J.F. Shogren, S.Y. Shin, and J.B. Kliebenstein. 1995. "Valuing Food Safety in Experimental Auction Markets." *American Journal of Agricultural Economics* 77(1):40–53.

Holt, C.A., and S.K. Laury. 2002. "Risk Aversion and Incentive Effects." *The American Economic Review* 92(5):1644–1655.

Kroll, S., T.L. Cherry, and J.F. Shogren. 2007. "Voting, Punishment, and Public Goods." *Economic Inquiry* 45(3):557–570.

Lusk, J.L., J.A. Fox, T.C. Schroeder, J. Mintert, and M. Koohmaraie. 2001. "In-Store Valuation of Steak Tenderness." *American Journal of Agricultural Economics* 83(3):539–550.

Lusk, J.L., F.B. Norwood, and J.R. Pruitt. 2006. “Consumer Demand for a Ban on Antibiotic Drug Use in Pork Production.” *American Journal of Agricultural Economics* 88(4):1015–1033.

Melkani, A., and N.M. Mason. 2018. “Tanzania Southern Highlands Farm Input Promotions-Africa Village-Based Agricultural Advisors Baseline Survey Report.” *Technical Report*:57.

Messer, K.D., H. Zarghamee, H.M. Kaiser, and W.D. Schulze. 2007. “New hope for the voluntary contributions mechanism: The effects of context.” *Journal of Public Economics* 91(9):1783–1799.

Palm-Forster, L.H., S.M. Swinton, F. Lupi, and R.S. Shupp. 2016. “Too Burdensome to Bid: Transaction Costs and Pay-for-Performance Conservation.” *American Journal of Agricultural Economics* 98(5):1314–1333.

Pardey, P.G., C. Chan-Kang, J.M. Beddow, and S.P. Dehmer. 2015. “Long-run and Global R&D Funding Trajectories: The U.S. Farm Bill in a Changing Context.” *American Journal of Agricultural Economics* 97(5):1312–1323.

Roosen, J., J.A. Fox, D.A. Hennessy, and A. Schreiber. 1998. “Consumers’ Valuation of Insecticide Use Restrictions: An Application to Apples.” *Journal of Agricultural and Resource Economics* 23(2):367–384.

Schechter, L. 2007. “Traditional trust measurement and the risk confound: An experiment in rural Paraguay.” *Journal of Economic Behavior & Organization* 62(2):272–292.

Shogren, J.F., S.Y. Shin, D.J. Hayes, and J.B. Kliebenstein. 1994. “Resolving Differences in Willingness to Pay and Willingness to Accept.” *The American Economic Review* 84(1):255–270.

Waldman, K.B., J.M. Kerr, and K.B. Isaacs. 2014. “Combining participatory crop trials and experimental auctions to estimate farmer preferences for improved common bean in Rwanda.” *Food Policy* 46:183–192.

Yackee, S.W. 2006. “Sweet-Talking the Fourth Branch: The Influence of Interest Group Comments on Federal Agency Rulemaking.” *Journal of Public Administration Research and Theory* 16(1):103–124.

1. COMMENTS, CONTRIBUTIONS, AND COMPLIANCE: EVIDENCE FROM A PUBLIC GOODS EXPERIMENT

1.1 Introduction

Designing regulatory policy for the governance of public goods presents two primary challenges for administrators: (i) how to motivate heterogenous individuals to pay for the provision of public goods (Burlando and Guala 2005; de Oliveira et al. 2015; Gallier et al. 2017; Ostrom 1990); and (ii) how to enforce compliance with rule-based contribution mechanisms (Anderson and Stafford 2003; Murphy and Stranlund 2007). While we regularly observe individuals voluntarily contributing to public goods in both lab (Ledyard 1995) and field (Andreoni 2006; Ostrom 1990) settings, there are many instances where low contributions and noncompliance jeopardize the provision of public goods. Monetary and non-monetary institutional mechanisms (Chaudhuri 2011; Fehr and Gächter 2000; Ostrom 1990) have been important tools used to foster public goods provision and to manage the commons.

Increasing attention has been devoted to the use of mechanisms not related to financial sanctions or rewards to motivate increased individual contributions to public goods (Chaudhuri 2011; Masclet et al. 2003; Rege and Telle 2004; Shang and Croson 2009). We investigate one pathway to motivate increased contributions and compliance with the provision of public goods: inviting individual participation in the rulemaking process. Individuals can participate in a rulemaking or policymaking process through a variety of means including voting, meeting with decisionmakers, attending open forums, or submitting comments to regulatory bodies. In this paper, we focus on comment-based participation which has long been a pillar of deliberative governance. Comments in this context are distinct from previously studied mechanisms because they allow for unstructured communication and are not associated with a fixed aggregation rule.

We design an experiment to identify the marginal impact of a comment mechanism on individual public goods provision. To do this, we implement a modified linear public goods game with two key treatments. First, we introduce a novel comment mechanism where individuals can provide unstructured feedback to influence a third-party rule-maker whose sole task is to determine a minimum contribution rule (MCR) at the beginning of the game. Second, we implement a costly probabilistic sanctioning mechanism for individuals who do not contribute at least the MCR to the public good. Combined, these treatments allow us to assess the impact of comments under a baseline voluntary contribution mechanism (VCM) as well as in the presence of complementary exogenous enforcement mechanisms.

Our results suggest that soliciting comments from individuals can help rule-makers overcome free-riding and the under-provision of public goods under certain institutional settings. First, we find little evidence of a direct effect of comments on individual contribution and compliance behavior. However, the comment mechanism significantly enhances contributions when implemented in the presence of the sanctioning mechanism, so that jointly, they yield the highest individual contributions to the public good. Second, we find that individual compliance with the MCR is also highest in the presence of comments paired with sanctioning.

We also find that comments affect the behavior of the rule-maker in this decision environment. Receiving comments increases the level of the MCR chosen by the rule-maker by 40% on average, independent of the exogenous sanctioning mechanism. This suggests that comments may have an indirect effect on players' contribution behavior via their effects on the MCR chosen by the rule-maker. Furthermore, we find that players respond to an increase in the MCR by increasing their contribution to the public good. We perform a mediation analysis and find that the effect of comments on the MCR explains approximately two-thirds of the increase in individual contributions to the public good.

1.2 Previous research

Our study is related to several lines of research related to public goods: voting, sanctioning, third-party public goods scenarios, and empirical research focused on comment-based participation. We highlight key studies in each of these areas.

1.2.1 Participation in public good decision making

Experimental research on the effects of individual participation in public good decision-making processes has been dominated by studies of voting rules and their effect on cooperation. There are two types of voting mechanisms of interest: binding and nonbinding. In a binding vote, the results obtained from some aggregation rule (e.g. majority rule) are enforced by an exogenous actor (Walker et al. 2000). Nonbinding votes, often observed in real-world assignment of contributions to public goods (e.g. environmental governing bodies and school boards), have no exogenous actor to enforce the outcome of a vote but are instead accompanied by the announcement of results (Kroll et al. 2007).

Evidence on the ability of voting to increase contributions to a public good is mixed. Binding majority-rule votes are found to induce more contributions to a public good than no-vote scenarios (Kroll et al. 2007; Walker et al. 2000). Walker et al. (2000) allow individuals to make proposals about contributions to the public good and then vote on adopting a rule via a majority rule or unanimity system in each round of the game. Subjects participating in both voting schemes increased contributions, even when a proposal was not adopted. Coordination in the face of a failed vote was likely due to the signaling system whereby all participants could see the proposals of other group members and gain information about their preferred rule. Kroll et al. (2007) follow a similar model and find that binding votes result in the highest contributions to the public good. Feld and Tyran (2002) show that voting on a punishment structure, rather than a level of public goods provision, also increases compliance. Focusing on common property goods, Gatiso and Vollan (2017) find that both voting for management rules and voting for leaders who choose management rules can increase cooperation and compliance over exogenous rule and leader determination.

On their own, nonbinding votes have little effect on contributions to public goods (Kroll et al. 2007; Messer et al. 2007). Kroll et al. (2007) allow individuals to make proposals about how much individuals should contribute to the public good, and then submit nonbinding votes for a preferred proposal. Despite no significant effect of nonbinding votes alone, they find that pairing nonbinding voting mechanisms with the opportunity to sanction free-riders sustains a higher level of cooperation than nonvoting mechanisms. Similarly, Messer et al. (2007) find that interacting nonbinding voting with cheap talk can reduce the rate of decay for contributions to the public good.

An important caveat to the voting literature is the existence of heterogeneous preferences among individuals that might impact the response to a voting rule. Specifically, we may be concerned that the individuals who vote in support of a rule are more likely to cooperate post-implementation than those opposed to a rule. Dal Bó et al. (2010) use a prisoner's dilemma experiment where individuals cast a majority vote about whether or not to impose fines on unilateral defectors. The vote is then randomly upheld or overruled. After controlling for selection effects, the authors find strong evidence that voting results in more cooperative behavior than exogenously imposed rules.

Another related line of work focuses on the use of rule-based contribution mechanisms where the lowest proposed public good level is selected and a burden-sharing rule is enforced (Dannenberg et al. 2014; Gallier et al. 2017; Kesternich et al. 2014, 2018; Orzen 2008). Although not a true voting process, the proposal stage of participation does induce engagement with rulemaking and establishes a defined outcome. Gallier et al. (2017) introduce voting mechanisms into this environment over the selection of the burden-sharing rule. Using both unanimity and majority rule procedures in the context of endowment heterogeneity, they show that if individuals can agree on a burden-sharing scheme then contributions to the public good increase. If individuals fail to reach a consensus via voting, however, there are lower contributions than in a baseline public goods game.

We contribute to this literature by implementing an open comment mechanism through which participants can try to influence a rule-maker's decision over a public goods contribution rule. Comments allow for less structure in participation than voting mechanisms, permitting the expression of significant amounts of heterogeneity in the decision-making process. Not only can an individual use a comment to propose a level of public good, but they can also use the opportunity to send a clearer signal about intention and rationale than is possible through a yes/no vote (Walker et al. 2000). The comments implemented in this study are differentiated from other forms of cheap talk (e.g. Bochet et al. 2006; Palfrey et al. 2017; Palfrey and Rosenthal 1991) because they are private in nature and only communicated to a third-party rule-maker. Private comments allow for participation in choosing the level of public goods provision without providing a mechanism for player communication before the VCM, thus isolating the impact of commenting from signaling or cheap talk.[1] This is different from other implementations of nonbinding voting

[1] Kroll et al. 2007 compare nonbinding voting to cheap talk. In their experiment, this is an accurate characterization because all players who must make a contribution decision are aware of the proposals of other group members and

where there may be significant effects of signaling, especially at the proposal stage for the contribution rule or the announcement of results (Kroll et al. 2007; Messer et al. 2007).

1.2.2 Experimental literature on sanctioning and enforcement

In addition to participation channels, sanctioning and enforcement can also induce greater contributions to a public good. The literature on sanctioning can broadly be divided into two categories: exogenous and endogenous. Anderson and Stafford (2003) conceptualize exogenous sanctions as being imposed by a third party and consisting of both the level of the financial penalty as well as a probability of being sanctioned. In this example, any resources kept in the private account were subject to sanctioning. The authors find that group contributions are increasing in the expected cost of punishment; however, the level of the sanction has a larger effect than the probability of being sanctioned. Qin and Wang (2013) investigate the expected costs of sanctioning where only the lowest contributor to the group account is sanctioned. While they lack a standard by which to measure compliance, they find that free riding in a public goods game is minimized when individuals face a 50% chance of losing their entire endowment.

Most experimental research on sanctioning mechanisms focus on the use of endogenous sanctioning mechanisms (e.g. Chaudhuri 2011; Fehr and Gächter 2000; Kroll et al. 2007; Tyran and Feld 2006). Seminal work by Fehr and Gächter (2000) illustrates how the opportunity to engage in costly punishment of free-riders significantly increases public goods provision. Other research has extended this literature in numerous directions including comparisons with non-monetary punishment (Masclet et al. 2003), the price-responsiveness of punishment (Anderson

---

have some idea of the division of votes (majority rule) before starting the game. Comments submitted publicly might have a similar effect.

and Putterman 2006), punishment efficiency (Nikiforakis and Normann 2008), and duration effects (Gächter et al. 2008).

Some studies allow for the selection of a sanctioning institution via voting (Dal Bó et al. 2010; Gürerk et al. 2006; Kamei et al. 2015; Kroll et al. 2007; Putterman et al. 2011; Sutter et al. 2010; Tyran and Feld 2006). Tyran and Feld (2006) find that mild sanctions are effective at increasing contributions to a public good when selected endogenously (via voting) but have no effect when imposed exogenously. Putterman et al. (2011) show that individuals who vote to design their own formal sanctioning scheme can converge on an efficient design within several rounds. Extending this work, Kamei et al. (2015) permit voting over the use of endogenous formal or informal sanctions. They find that while costless formal sanctions are popular and efficient, the introduction of costs shifts voter preferences towards the use of informal sanctions.

We specify an exogenous probabilistic sanctioning mechanism in our experimental design where under-contribution to the public good relative to a specified rule exposes an individual to a positive probability of sanctioning after each contribution decision. This is similar to Anderson and Stafford (2003) in the use of exogenous enforcement but differs in two key ways. First, our standard for measuring compliance is a player's contribution to the public good vis-à-vis the MCR determined by a third-party rule-maker. In fact, this may more closely represent a typical regulatory standard compliance setting. Second, the sanction for a violation is fixed, and does not vary with the magnitude of the violation. For example, an individual who under-contributes to the public good by one unit faces the same expected cost of sanctioning as an individual who under-contributes by five units. This assumption simplifies the calculation for noncompliance – we would expect any individual who intends to engage in free-riding to contribute zero resources towards the public good. We implemented this mechanism to capture the idea that there is often a minimum penalty assessed for any rule violation. Additionally, regardless of the level of the violation, there is likely to be a fixed cost associated with coming into noncompliance. Our sanctioning mechanism captures both of these effects by imposing a strong, fixed penalty on noncompliers with a positive probability.

1.2.3 Third-party actors in an experimental setting

Another area of public goods-related research deals with individual interaction with a third-party actor. Most studies in this literature focus on sanctioning and employ a third party who is able to observe individual behavior and choose both who is sanctioned and at what level (Bernhard et al. 2006; Carpenter and Matthews 2012; Fehr and Fischbacher 2004; Henrich et al. 2006; Kosfeld and Rustagi 2015). This model approximates the behavior of many types of regulatory environments governing public goods whereby an external authority or agency chooses when to sanction individuals for noncompliance or under-contribution (e.g. the Environmental Protection Agency (EPA) choosing to prosecute a Clean Water Act violation).

We implement comments in a unique lobbying environment combining elements of endogenous and exogenous rule making. While individuals are invited to provide input on public goods provision via comments, ultimate authority rests with an external player who is only tasked with choosing the MCR. This structure provides an exogenous decision-making authority that may or may not be influenced by comments. This bears significant resemblance to a third-party punishment game; however, instead of sanctioning, the third party sets the level of public goods provision. Players have incentives to try to leverage their comment to garner influence with the decision-making authority. The rule-maker is compensated only for choosing the rule and not for the level of public goods provided to eliminate any incentives to choose a high rule. This better describes the environment individuals might face when engaging with real world regulatory participation mechanisms. For example, rule-makers at the EPA receive the same salary regardless of the level or number of pesticide regulations issued (Cropper et al. 1992). Additionally, this element of the design allows us to analyze the impact of comment-based participation on the third-party actor and consider the direct and indirect effects of comment on both rule choice and individual contribution and compliance behavior. Environments without a third-party player cannot make this distinction.

1.2.4 Empirical literature on comments in rulemaking process

Finally, our experimental analysis provides a complement to the large empirical literature on the impacts of participation mechanisms and decision rules on economic behavior and outcomes (Bardhan 2000; Black and Lynch 2001; Feld and Frey 2007). A smaller literature focuses on comments and regulatory outcomes specifically. Magat et al. (1986) examine how comments influence best practical control technology effluent regulations at the EPA. They find that the number of industry comments received, and the relative strength of regulated industries has no impact on the final rules generated by policymakers. Conversely, Cropper et al. (1992) find strong support for the power of participation in EPA pesticide regulations. Comments from grower organizations reduce the probability that cancer-causing pesticides are banned while comments from environmental groups result in stricter regulatory controls. Similarly, Yackee (2006) finds strong evidence that bureaucratic rule-makers respond to comments by changing final regulations to more closely match the ideal points of the interest groups that submit comments. The challenge in this literature is the underlying heterogeneity present in the salience of the issues being discussed, the types of regulations under consideration, and the linkages of comments to the timeline of rulemaking decisions (Crow et al. 2016). Our study contributes to this literature by designing an experimental protocol that is able to capture both the effect of participation on regulatory outcomes and ex-post individual compliance.

1.3 Behavior in the public goods game

We embed our design situation in a standard linear public goods game with a VCM (e.g. Andreoni 1995; Isaac and Walker 1988; Ledyard 1995). In the standard VCM, each individual i in a group of size N maximizes her utility U_i by choosing what portion of an endowment she would like to allocate to a private account versus to a group account.

$$U_i = \left(E_i - x_i\right) + P\sum_{i \in N} x_i \tag{1}$$

Equation (1) models individual utility in this game. E_i is the player's initial endowment (in points) at the beginning of the game. Let x_i be the number of points the individual chooses to contribute to the group account. Individual earnings from the private account are then given by $\left(E_i - x_i\right)$, where individuals get the full value of any points placed in their private account. Individual earnings from the group account are calculated as a multiple of total group contributions, where P is the marginal per capita return (MPCR). The MPCR is the marginal return to an individual of a one-point contribution to the public good. Thus, an individual's total earnings from the group account are given by $P\sum_{i \in N} x_i$. Given a group size of N, each point placed in the group account returns NP points to the group as a whole.

There exist rates of return that may make it Pareto optimal for all individuals to allocate their entire endowment to the group account while at the same time incentivizing individuals to only contribute to their private account. These incentives are described in Croson (2007), where the pure public goods problem arises when $\frac{1}{N} < P < 1$. In this environment, classical economic theory predicts an optimal contribution of $x_i^* = 0$, which is also the best response to the allocation decisions of any other member of the group (i.e. $\frac{\partial x_i^*}{\partial x_j} = 0 \; \forall i \neq j \in N$). In this case, we would expect a unique equilibrium where all participants contribute zero resources to the public good and would have earnings equal to their initial endowment.

Adding an exogenous probabilistic enforcement mechanism to the linear public goods environment changes the nature of equilibrium play. We implement enforcement as a function of both the level of the fine and the probability of being caught in order to better reflect the complexity of most regulatory monitoring schemes (Anderson and Stafford 2003; Qin and Wang 2013). Suppose there is a rule, r, that specifies the minimum number of points that an individual must contribute to the group account. Participants in a state of noncompliance with the rule ($x_i < r$) will be observed (caught) with probability λ and will face a fine of q points. Expected returns for the individual now become:

$$U_i = \begin{cases} \left(E_i - x_i\right) + P\sum_{i \in N} x_i, \text{ if } x_i \geq r \\ \left(E_i - x_i\right) + P\sum_{i \in N} x_i - \lambda q, \text{ if } x_i < r \end{cases} \tag{2}$$

The decision faced by the individual is now a threshold problem. Building on the incentives provided when $\frac{1}{N} < P < 1$, the contribution decision to maximize individual payoffs now becomes a choice between contributing zero points to the group account or setting $x_i = r$. Comparing the expected payoffs under each scenario, the optimal decision rule is:

$$x_i^* = \begin{cases} r \text{ if } r < \dfrac{\lambda q}{1-P} \\ 0 \text{ otherwise} \end{cases} \tag{3}$$

That is, the optimal response to a contribution rule in the presence of an enforcement mechanism is determined by the MPCR (P) as well as the probability of being caught in a state of noncompliance (λ) and the severity of the fine (q). With this enforcement scheme, standard theory would now predict that we may observe positive contributions to the group account in equilibrium depending upon the parameterization of the problem.

The addition of a comment mechanism may have further impacts on participant behavior. One theoretical link between comments and behavior is built on a view of procedural participation as a means to legitimize final rules and outcomes (Bouma et al. 2014; Bouma and Ansink 2013; Feld and Tyran 2002; Tyran and Feld 2006; Verba et al. 1978). When individuals are engaged in the process or procedures of policymaking, they are more likely to view the outcomes of that process as fair (Tyler 1990), even when outcomes do not reflect an individual's ideal policy (Feld and Tyran 2002; Thibaut and Walker 1975). Because comments are a form of direct engagement with policymakers, we would expect individuals who are exposed to comment mechanisms to be more likely to view a rule or regulation as legitimate or reasonable and therefore more likely to contribute to a public good and/or comply with a contribution rule. Dal Bó et al. (2010) show, using a laboratory experiment, that the effects of participation via voting on individual behavior are independent of any information transmission that might occur between subjects. This suggests that it is participation, and not a learning or signaling effect, that may drive increased contributions to the public good. Furthermore, the legitimacy of rules and outcomes can be bolstered when there exist complementary institutions to enforce compliance/sanction non-compliance (Kroll et al.

2007). While these enforcement mechanisms might not be enough to independently ensure compliance, they may promote an increased sense of legitimacy in that all individuals are expected to adhere to the requirements of a regulation (Tyler 1990).[2]

A second mechanism through which comment-based participation might influence contribution behavior and compliance is through individual learning about the policymaking process and the need for a rules-based solution to a problem. When individuals are invited to comment on a proposed regulation or policy, they might be more inclined to invest effort in understanding the rule and what impact it may have on them and others, which could lead to higher rates of compliance if the rule is implemented. While this mechanism is unlikely to be salient when considering relatively straightforward regulations, it is important to consider when thinking about compliance in more complex regulatory regimes. At the same time, their investment in understanding a rule and providing feedback may also cause a rule-maker to choose a rule closer to an individual's preferred rule. In this setting, comments help rule-makers learn about the preferences of participants, and in the public good environment, all players are likely to benefit from a higher rule.

Based on previous research on participation and the economic model underpinning our public goods game, this paper tests four main hypotheses: **(H1)** Comments will increase individual contributions to the public good. **(H2)** Comments will increase compliance with the MCR chosen by the rule-maker. **(H3)** The positive effect of comments on compliance will be larger when the expected cost of noncompliance is higher (e.g., in the presence of a probabilistic enforcement mechanism). **(H4)** Rule-makers receiving comments from players will choose a higher MCR.

---

[2] See Bouma and Ansink (2013) for a full discussion of how notions of legitimacy may influence individual behavior and preferences over the management of common pool resources.

1.4 Experimental design

The experiment follows a 2X2 design (see Table 1.1) involving participation and enforcement and results in three treatment groups and a control group (with no participation or enforcement). Participation involved either no participation or the opportunity for comments. In treatments with no participation, a randomly chosen rule-maker (explained further below) chooses a non-binding MCR (i.e., minimum number of points participants should contribute to a group account) with no comments or input from other players. In the comment treatments, players are given the opportunity to submit private comments on their preferences for the non-binding MCR to the rule-maker, who is given an opportunity to read the comments before choosing the MCR. Note that there is no restriction on choice of MCR in that it can be any whole number ranging from zero to the entire endowment. Treatments with enforcement implement a probabilistic punishment mechanism (described below) that, with a given probability, implements a fine if the individual's contribution is less than the MCR.

Table 1.1: Experimental treatments

Participation Treatment	**Enforcement Treatment**	
	No Enforcement	Enforcement
No Participation	(Control) 55 Participants	(Treatment 2) 50 Participants
Comment	(Treatment 1) 55 Participants	(Treatment 3) 55 Participants

The experiment was implemented using oTree software (Chen et al. 2016) and run with students in a computer lab at Michigan State University in April 2016 and February-March 2017. Students were recruited from a sample of all undergraduates enrolled at Michigan State University using ORSEE (Greiner 2015). Treatments were conducted in 12 sessions, with three sessions for each treatment. In total, there were 215 participants and the number of participants per treatment group varied from 50 to 55 (see Table 1.1). Following Sefton et al. (2007), all sessions included two stages. Stage 1 involved five rounds of a standard VCM game to acquaint participants with the play of the game. Stage 2 consisted of five separate games (with five rounds per game; 25 total rounds) that included the control/treatment conditions. Both Stage 1 and 2 VCM games were played in groups of five.[3] Subjects were not permitted to communicate with one another in any way other than through their decisions in the experiment. At the end of the experiment, subjects were paid their earnings from their decisions in private and in cash.

The Stage 2 games involved two participant roles – a rule-maker and regular players. Each group of five consisted of four regular players (heretofore referred to as players), and a rule-maker. A given participant's role remained the same for all five rounds of a given game, but the experiment was structured so that each participant had the opportunity to be the rule-maker during one of the five games. Specifically, each subject was randomly assigned to be a rule-maker for one of the games and, for each game, the remaining participants (i.e., the players) were randomly sorted into new groups to minimize potential social effects across games.

Parameters were chosen to be largely consistent with previous public goods experiments. All players were provided with an endowment of 25 points (Croson 2000; Crumpler and Grossman 2008; Ostrom et al. 1992) and asked to divide those points between a group account and a private account. Players kept the full value of any points allocated to the private account. The MPCR was set at 0.4 such that incentives would be consistent with a pure public goods problem (Croson and

[3] Because the experiment was designed around groups of five individuals, each session was run with 15-20 subjects. The enforcement only treatment had an additional session where several registered participants did not show up to the experiment and the session was conducted with 15 subjects.

Marks 2000; Fehr and Gächter 2000). With an MPCR of 0.4, each player in the group received 0.4 points for each point allocated to the group account. This meant that each point placed in the group account returned 1.6 points to the group of four players as a whole at the end of a round.

We implement our comment process using what is, to the best of our knowledge, a novel variant of third-party involvement in a public goods game. While most studies have used the existence of a third party to determine and enforce sanctions (Charness et al. 2008; Fehr and Fischbacher 2004; Fiedler and Haruvy 2017; Kosfeld and Rustagi 2015), we use the third-party to set the MCR for the public good. In comment treatments, before the first round of play, all players were provided the opportunity to send the rule-maker comments regarding their preferences for the choice of MCR. After all players submitted their comments, these comments were provided to the group's rule-maker anonymously and in a random order. The rule-maker was then given an opportunity to read the comments and choose an MCR. This MCR was then used for all five rounds of the game. For completing the task of reading the comments and choosing an MCR, the rule-maker was paid a salary of 25 points per round (125 points total). The no comment treatments differed only in that the rule-maker in each group chose an MCR without receiving any comments from players. Once the MCR was chosen, the rule was communicated to players who proceeded with the public goods VCM.

In enforcement treatments, subjects were informed in the instructions about the existence of an exogenous punishment mechanism. The probability of being punished in any given round was set at 50% and the fine was equal to the initial endowment, 25 points. This sanctioning mechanism was chosen to determine if relatively robust findings on the effects of sanctions would drown out any effect of comments (Anderson and Stafford 2003; Qin and Wang 2013). At the end of each round, the software identified each individual who had contributed less than the MCR to the group account and determined whether or not they were to be sanctioned according to the parameters described above.

At the end of each round, the players viewed a summary screen with the following information: (i) the number of points they contributed to their private and group accounts; (ii) the MCR chosen by the rule-maker; (iii) whether or not they were sanctioned, and the level of the sanction (enforcement-related treatments only); (iv) the anonymous contributions of other group members to the group account; and (v) their total earnings from the current round. After each five-round game, participants were randomized into new groups and roles as described above and the process was repeated. Point earnings were totaled across all five games and converted to dollars at a rate of $0.03 per point. At the end of the experiment, subjects were presented with a performance summary describing their earnings in each game and their total earnings for the overall experiment.[4]

1.5 Results

Each of the 215 participants made contribution decisions to a private and group account during four of the five games. Given five rounds in each game, this results in a dataset of 4,300 contribution decisions pooled across all games, rounds, and treatments. Each participant was also randomly assigned to the role of rule-maker during one of the five games. This means that we observe 215 decisions over the MCR. In the following sections, we will explore four key outcomes of interest in this experiment: each player's contributions to the group account (public good), two measures of non-compliance with the MCR (binary and relative), and the level of the MCR.[5]

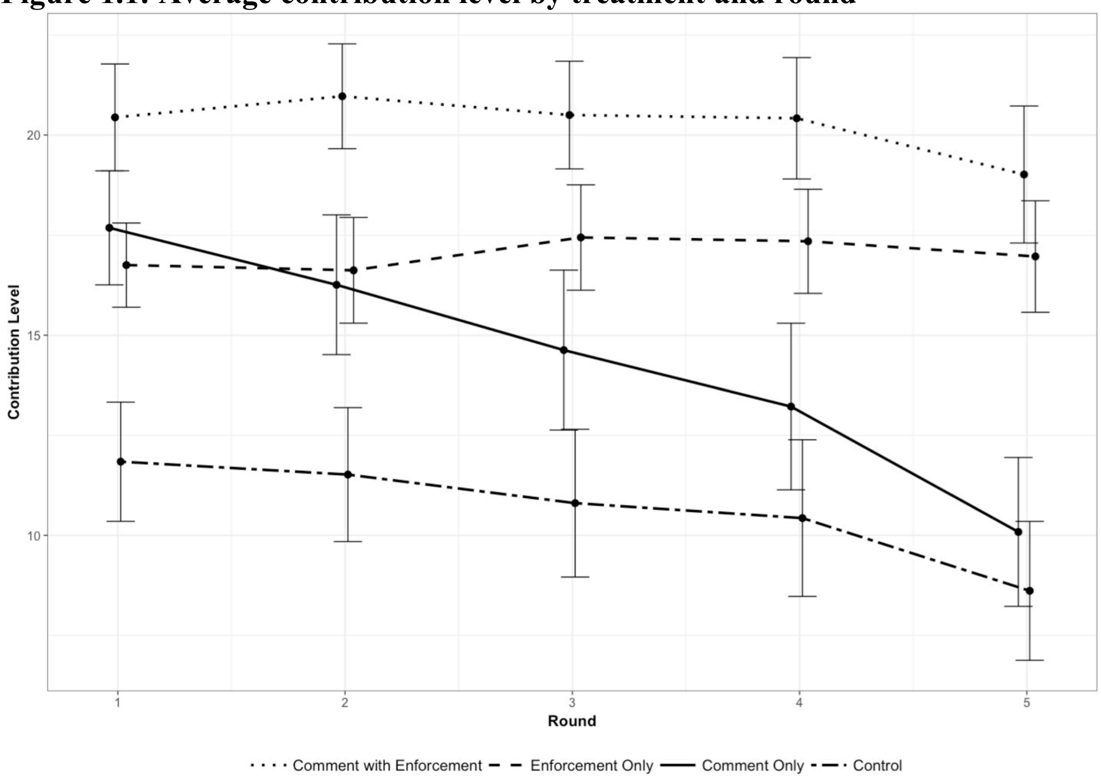
---

[4] Subjects were also compensated with a $10 participation fee for showing up to their assigned experimental session.
[5] Appendix Table 1A.1 presents summary statistics for the key outcomes of interest in this experiment. Panel A presents averages pooled across all rounds and games and Panel B presents p-values from pairwise mean comparisons. The results are similar when compared via nonparametric Mann-Whitney U tests.

1.5.1 Contributions to the public good

We first examine the effects of participation and enforcement treatments on individual contributions to the public good. Figure 1.1 plots, by round, the average individual contribution level to the public good for each treatment. Note that averages are taken across all groups in the experiment such that the round one average includes contributions from round one in all five games. Standard errors are calculated at the group level.

Figure 1.1: Average contribution level by treatment and round



Note: Standard errors calculated around group means.

Immediately, we see that all treatments appear to incentivize higher levels of contributions to the public good than the control group. The comment with enforcement treatment induces the highest average level of contributions across all rounds. Subjects exposed to the enforcement only treatment exhibit slightly lower contribution levels than those in the comment with enforcement group. Both treatments are able to maintain those contribution levels across rounds on average. In contrast, the comment only treatment starts out with high levels of contributions (similar to enforcement only), but on average, decays over rounds to a contribution level similar to the control group. Plotting these simple treatment averages suggests that comments alone may not be able to sustain increased levels of contributions to a public good over the long-term.

We now turn to regression analysis to quantify the effects of the different treatments on a player's contributions to the public good (group account). To account for corner solutions in subject contribution decisions (many players contributed zero points to the group account), we use a Tobit model where the dependent variable is the number of points contributed to the group account. Independent variables include the MCR chosen by the rule-maker, indicator variables for each treatment with the control group serving as the reference category, round dummy variables, group dummy variables, and interaction variables between treatment and round as well as treatment and game. Table 1.2 reports the average partial effects (APEs) from our Tobit specifications as well as the results of hypothesis tests comparing treatment categories. Coefficients from the Tobit regressions are presented in Appendix Table 1A.2.[6] [7]

---

[6] OLS results are consistent with the Tobit results presented here.

[7] We report the APEs as our main results for two reasons. First, the magnitude of an APE is directly interpretable as the estimated average change in the dependent variable given a one unit increase in the independent variable of interest in the continuous independent variable case and given a discrete change in the independent variable of interest in the binary independent variable case, holding other factors constant. In a model without interaction of squared terms, the coefficients from nonlinear-in-parameters models like Tobit models (and the probit and fractional response probit models estimated below) indicate the *direction* of the effects of the independent variables (positive or negative) but must be scaled in order for their *magnitudes* to be interpretable. Second, in some of our model specifications we include interactions between the treatment indicator variables and round and/or game dummies. APEs give an overall summary measure of the effect of a given independent variable on the dependent variable taking into account these interaction terms (and squared terms, if present). See Wooldridge (2010) for details.

Table 1.2: APEs of treatment on contributions to the public good (Tobit estimates)

	(1)	(2)	(3)	(4)	(5)
Minimum Contribution Rule	0.420***	0.420***	0.419***	0.446***	0.445***
	(0.041)	(0.041)	(0.041)	(0.041)	(0.042)
Comment Only (C)	-0.387	-0.381	-0.381	-0.593	-0.642
	(1.107)	(1.107)	(1.101)	(1.092)	(1.068)
Comment with Enforcement (CE)	5.990***	5.978***	5.994***	5.752***	5.758***
	(0.962)	(0.962)	(0.967)	(0.961)	(0.936)
Enforcement Only (E)	3.134***	3.131***	3.166***	3.027***	3.064***
	(0.860)	(0.860)	(0.866)	(0.854)	(0.857)
Round 2		-0.330	-0.386*	-0.326	-0.364*
		(0.204)	(0.203)	(0.204)	(0.203)
Round 3		-0.897***	-0.975***	-0.892***	-0.958***
		(0.267)	(0.263)	(0.267)	(0.261)
Round 4		-1.232***	-1.348***	-1.229***	-1.339***
		(0.303)	(0.291)	(0.303)	(0.290)
Round 5		-2.957***	-3.097***	-2.955***	-3.094***
		(0.365)	(0.322)	(0.364)	(0.319)
Game 2				0.179	0.200
				(0.829)	(0.792)
Game 3				-1.295	-1.498
				(0.945)	(0.928)
Game 4				-0.889	-1.078
				(0.971)	(0.953)
Game 5				-1.396	-1.579*
				(0.944)	(0.928)
Treatment x Round Interaction	No	No	Yes	No	Yes
Treatment x Game Interaction	No	No	No	No	Yes
Observations	4300	4300	4300	4300	4300
Hypotheses			**p-value**		
H_0: C = CE	0.000	0.000	0.000	0.000	0.000
H_0: C = E	0.000	0.000	0.000	0.000	0.000
H_0: CE = E	0.000	0.000	0.000	0.000	0.000

Note: Dependent variable is the number of points contributed to the group account from zero to 25. Standard errors in parentheses and clustered at the group level. Significance denoted by *** p<0.01, ** p<0.05, * p<0.1 respectively.

We find strong evidence that comments increase contributions to the group account, but only in the presence of an enforcement mechanism. Focusing on the results in model (1), we observe no effect of the comment treatment on individual contributions to the public good, when compared to the control group. However, the comment with enforcement treatment exhibits a strong effect on contribution behavior, resulting in a significant increase in individual contributions of 5.99 points on average, or 24% of the initial 25-point endowment. This is a larger effect than we find with enforcement only, where contributions increase by 3.13 points on average, equivalent to 13% of the initial endowment. The difference between the comment with enforcement and the enforcement only treatments is statistically significant.

Contrary to the initial hypothesis **H1**, these findings suggest that in isolation, an invitation to provide comments to a neutral or bureaucratic third-party rule-maker is not sufficient to induce higher individual contributions to a pure public good in a VCM environment. Instead, when comments are combined with a complementary enforcement mechanism, realized contribution levels exceed those of the enforcement only treatment. This validates our expectations in hypothesis **H3** that comments provide stronger incentives for contributions and compliance in the presence of probabilistic punishment. Thus, in environments where enforcement mechanisms are already in place, including more individuals in deliberative decision-making through comments may represent a pathway to achieving closer to Pareto optimal outcomes and increased provision of public goods. One reason for these findings could be that exogenous enforcement mechanisms are replacing private retaliation on the part of players. Players who participate in choosing an MCR via comment may have an expectation of cooperation by other group members. If cooperation is not observed from another player, individuals may be more likely to withhold contributions from the group account as a form of private punishment for noncooperation. When an exogenous punishment mechanism is present, individuals may feel exogenous sanctions are enough to punish noncompliance and there is no need to resort to private punishment via decreased allocation to the group account.

Model (2) adds to the analysis controls for the round in which a contribution decision occurred but finds little change in the relationship between our experimental treatments and contribution behavior. Consistent with previous research (Ledyard 1995), we observe that beginning with round 3, contributions to the group account decline as the game progresses. The largest decline in cooperation occurs in round 5 where contributions to the group account fall by 2.96 points on average relative to the first round, *ceteris paribus*. This decline is expected because subjects know they will not interact with the same group again, so they are insulated from any retaliatory behavior from a low contribution decision. Model (3) allows for heterogeneous treatment effects over rounds but the results are unchanged. Finally, model (4) controls for the game and round in which a contribution decision was made while model (5) includes interactions between treatment and both round and game. Our main results are robust to these alternative specifications.

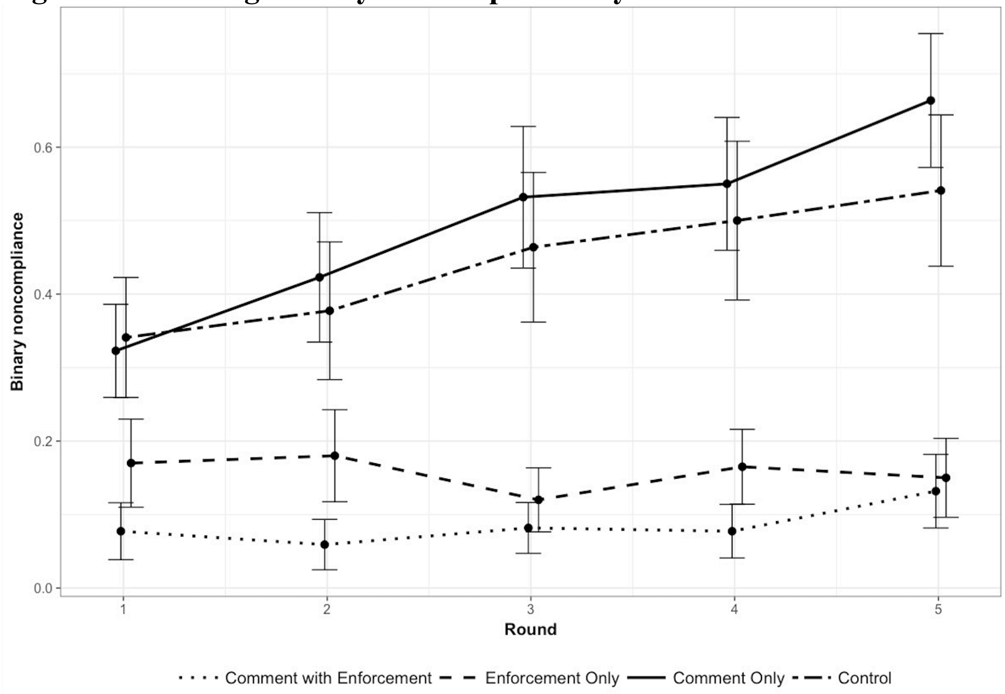
Across all models, we find strong evidence that a higher MCR induces individuals to contribute more points to the public good (group account). In model (1), a one-point increase in the MCR raises individual contributions by 0.42 points on average. Intuitively, this makes sense because individuals respond to the pro-public good choices of a third-party by also raising their contributions. However, subjects clearly do not respond perfectly to an increase in the MCR (i.e., 0.42 is well below one).

1.5.2 Comments and compliance with MCR

We next consider how the comment only, enforcement only, and comment with enforcement treatments impact compliance with the MCR chosen by the policymaker. We consider two measures of noncompliance: binary and relative. Binary noncompliance is simply whether or not an individual complies with the MCR. This measure is coded as a zero if the individual contribution to the group account is greater than or equal to the MCR, and a one if the individual contribution is less than the MCR. Relative noncompliance is the proportion of the MCR not accounted for by the individual allocation to the group account. For example, if the MCR is 10 points, an individual who contributes 3 points will have a relative noncompliance measure of 0.7. Higher values indicate that the individual contribution is further away from the MCR. A score of 1 implies no points were contributed to the group account (and the MCR was nonzero), whereas a score of zero indicates that the player's allocation to the group account was greater than or equal to the MCR.

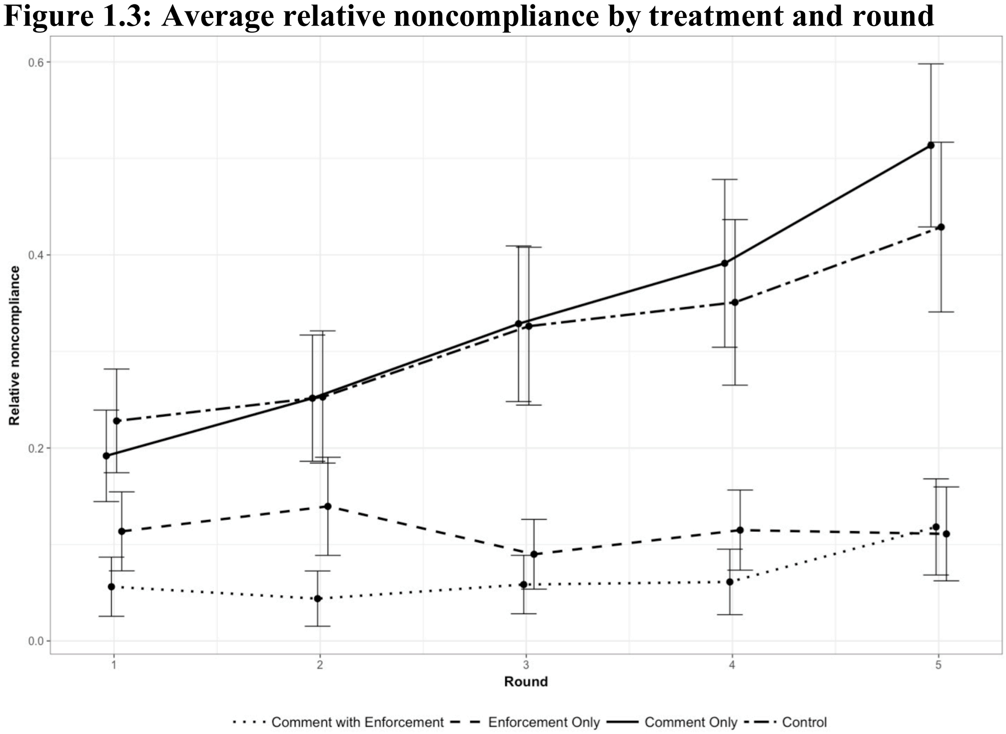
We begin by looking at plots of average binary noncompliance and relative noncompliance by round in Figures 1.2 and 1.3, respectively. In both figures, similar patterns for noncompliance emerge. We see very little difference in average noncompliance rates between the control group and the comment only treatment, with both trending towards higher rates of noncompliance across rounds. Additionally, we see that the comment with enforcement treatment generally appears to minimize both measures of noncompliance; however, the enforcement only treatment exhibits similar performance.

Figure 1.2: Average binary noncompliance by treatment and round



Note: Standard errors calculated around group means.

Figure 1.3: Average relative noncompliance by treatment and round



Note: Standard errors calculated around group means.

To estimate the effect of treatment on binary noncompliance, we estimate probit models with APEs reported in Table 1.3.[8] (Coefficients from the probit regressions are presented in Appendix Table 1A.3.) The control group again serves as the excluded reference category. Here we include the same set of independent variables that were included in the contribution analysis, except we exclude the MCR as a regressor due to its use in the construction of the dependent variables measuring noncompliance.

Focusing on the model (1) results presented in Table 1.3, we find no evidence that the comment only treatment has an effect on individual compliance with the MCR when compared to the control group. Refuting **H2**, comments appear to be insufficient to induce higher compliance with the chosen MCR – a finding consistent with the lack of change in individual contribution levels. Our estimates suggest that the enforcement only treatment significantly decreases the probability that an individual engages in noncompliance with the MCR by 28.8 percentage points on average, while the comment with enforcement treatment significantly reduces it by 35.9 percentage points on average. Additionally, we find that the APE for the comment with enforcement treatment is significantly larger (in absolute value) than the enforcement only treatment, again suggesting an additional benefit that can be gained when including comments in the presence of enforcement regimes. These findings are robust to the inclusion of round dummies, game dummies, and their respective interactions with the experimental treatments as illustrated in models (2) through (5).

[8] OLS results are consistent with the probit and fractional response results presented here.

Table 1.3: APEs of treatment on binary noncompliance (probit estimates)

	(1)	(2)	(3)	(4)	(5)
Comment Only (C)	0.054	0.054	0.054	0.054	0.054
	(0.059)	(0.059)	(0.059)	(0.057)	(0.056)
Comment with Enforcement (CE)	-0.359***	-0.359***	-0.359***	-0.357***	-0.359***
	(0.048)	(0.048)	(0.048)	(0.046)	(0.046)
Enforcement Only (E)	-0.288***	-0.286***	-0.288***	-0.285***	-0.287***
	(0.049)	(0.049)	(0.049)	(0.048)	(0.048)
Round 2		0.029**	0.033**	0.029**	0.032**
		(0.014)	(0.013)	(0.014)	(0.013)
Round 3		0.068***	0.074***	0.068***	0.074***
		(0.015)	(0.014)	(0.015)	(0.014)
Round 4		0.091***	0.098***	0.091***	0.097***
		(0.015)	(0.014)	(0.015)	(0.014)
Round 5		0.142***	0.148***	0.142***	0.148***
		(0.018)	(0.016)	(0.018)	(0.016)
Game 2				0.003	0.004
				(0.042)	(0.042)
Game 3				0.111**	0.114**
				(0.050)	(0.051)
Game 4				0.100**	0.105**
				(0.046)	(0.047)
Game 5				0.130***	0.136***
				(0.048)	(0.048)
Treatment x Round Interaction	No	No	Yes	No	Yes
Treatment x Game Interaction	No	No	No	No	Yes
Observations	4300	4300	4300	4300	4300
Hypotheses			**p-value**		
H_0: C = CE	0.000	0.000	0.000	0.000	0.000
H_0: C = E	0.000	0.000	0.000	0.000	0.000
H_0: CE = E	0.003	0.003	0.003	0.003	0.002

Note: Dependent variable is equal to one if the subject does not contribute at least the MCR to the group account and is zero otherwise. Standard errors in parentheses and clustered at the group level. Significance denoted by *** p<0.01, ** p<0.05, * p<0.1 respectively.

Table 1.4: APEs of treatment on relative noncompliance (fractional response estimates)

	(1)	(2)	(3)	(4)	(5)
Comment Only (C)	0.018	0.017	0.018	0.017	0.018
	(0.048)	(0.048)	(0.048)	(0.046)	(0.046)
Comment with Enforcement (CE)	-0.250***	-0.250***	-0.250***	-0.249***	-0.250***
	(0.037)	(0.037)	(0.037)	(0.037)	(0.036)
Enforcement Only (E)	-0.204***	-0.201***	-0.204***	-0.202***	-0.204***
	(0.038)	(0.038)	(0.038)	(0.037)	(0.037)
Round 2		0.023**	0.024***	0.023**	0.024***
		(0.010)	(0.009)	(0.010)	(0.009)
Round 3		0.051***	0.055***	0.051***	0.055***
		(0.012)	(0.011)	(0.011)	(0.011)
Round 4		0.079***	0.084***	0.079***	0.084***
		(0.012)	(0.012)	(0.012)	(0.012)
Round 5		0.144***	0.149***	0.145***	0.149***
		(0.015)	(0.014)	(0.015)	(0.014)
Game 2				0.014	0.015
				(0.033)	(0.033)
Game 3				0.094**	0.096**
				(0.039)	(0.039)
Game 4				0.085**	0.089**
				(0.040)	(0.041)
Game 5				0.095**	0.098**
				(0.038)	(0.038)
Treatment x Round Interaction	No	No	Yes	No	Yes
Treatment x Game Interaction	No	No	No	No	Yes
Observations	4300	4300	4300	4300	4300
Hypotheses			**p-value**		
H_0: C = CE	0.000	0.000	0.000	0.000	0.000
H_0: C = E	0.000	0.000	0.000	0.000	0.000
H_0: CE = E	0.024	0.019	0.024	0.020	0.019

Note: Dependent variable is equal to the proportion of the MCR not accounted for by the individual contribution to the group account. Standard errors in parentheses and clustered at the group level. Significance denoted by *** p<0.01, ** p<0.05, * p<0.1 respectively.

Fractional response probit APEs for the effect of treatment on relative noncompliance are presented in Table 1.4 and the associated coefficient estimates are provided in Appendix Table 1A.4.[9] The findings for the effect of treatment on relative noncompliance are similar to those of binary noncompliance. Highlighting the results in Table 1.4 model (1), we find that the enforcement only treatment significantly reduces relative noncompliance by 20.4 percentage points on average, and the comment with enforcement treatment reduces it by 25.0 percentage points on average. Comments alone have no discernable effect on relative noncompliance when compared to the control group. Coupled with the results above on binary noncompliance, these findings suggest that comment mechanisms in the presence of enforcement reduce both the *probability* that an individual engages in noncompliance and the *extent* to which individuals deviate from the chosen MCR, on average while holding other factors constant. Additionally, comment with enforcement mechanisms reduces relative noncompliance by significantly more than enforcement alone. Thus, even for individuals who choose not to contribute the full amount of the MCR, participation via comment may encourage behavior that is closer to a state of compliance, conditional on enforcement.

In the regression analyses for both noncompliance measures (binary and relative), we observe significant increases in the probability and extent of noncompliance during later games. These can be seen in the positive and significant APEs on the game indicators in models (4) and (5) of Tables 1.3 and 1.4. There are two possible explanations for this observed behavior. First, it could be that as the experiment progresses from game to game, individuals might gradually become less cooperative and reduce contributions to the public good. While within a game we observe decreases in contributions (and thus increases in noncompliance across rounds), this

---

[9] Recall that relative noncompliance is a proportion between zero and one, hence the use of a fractional response probit model. See Papke and Wooldridge (1996) for details.

explanation is inconsistent with the results presented in Table 1.2 where the game has no significant effect on contribution behavior. The second explanation is that rule-makers might choose higher minimum contribution rules in the later games of the experiment. If this were the case, then the noncompliance measures would increase from game to game because (as shown in Table 1.2) there is a less than one to one correspondence between individual contributions to the public good and any increases in the MCR. As the discussion of rule-maker behavior below shows, the latter explanation is more likely.

1.5.3 Comments and rule-maker behavior

Our use of third-party rule-maker in the experiment also allows for an analysis of how the various treatments might affect the behavior of the rule-maker when choosing a contribution rule. In general, we see substantial dispersion in the contribution rules chosen in this experiment. Figure 1.4 presents a histogram of the contribution rules identified by rule-makers and broken down by treatment. As noted earlier, the rule-maker can choose any whole number between zero and 25 points for the MCR. The most frequent MCR chosen was 25 points, with spikes in frequency at five-point increments. Given this variation in rule choice, regression analysis can approximate the effect that exposure to comments (with and without enforcement) had on the MCR selected by a policymaker.

Figure 1.4: Histogram of MCR chosen by treatment

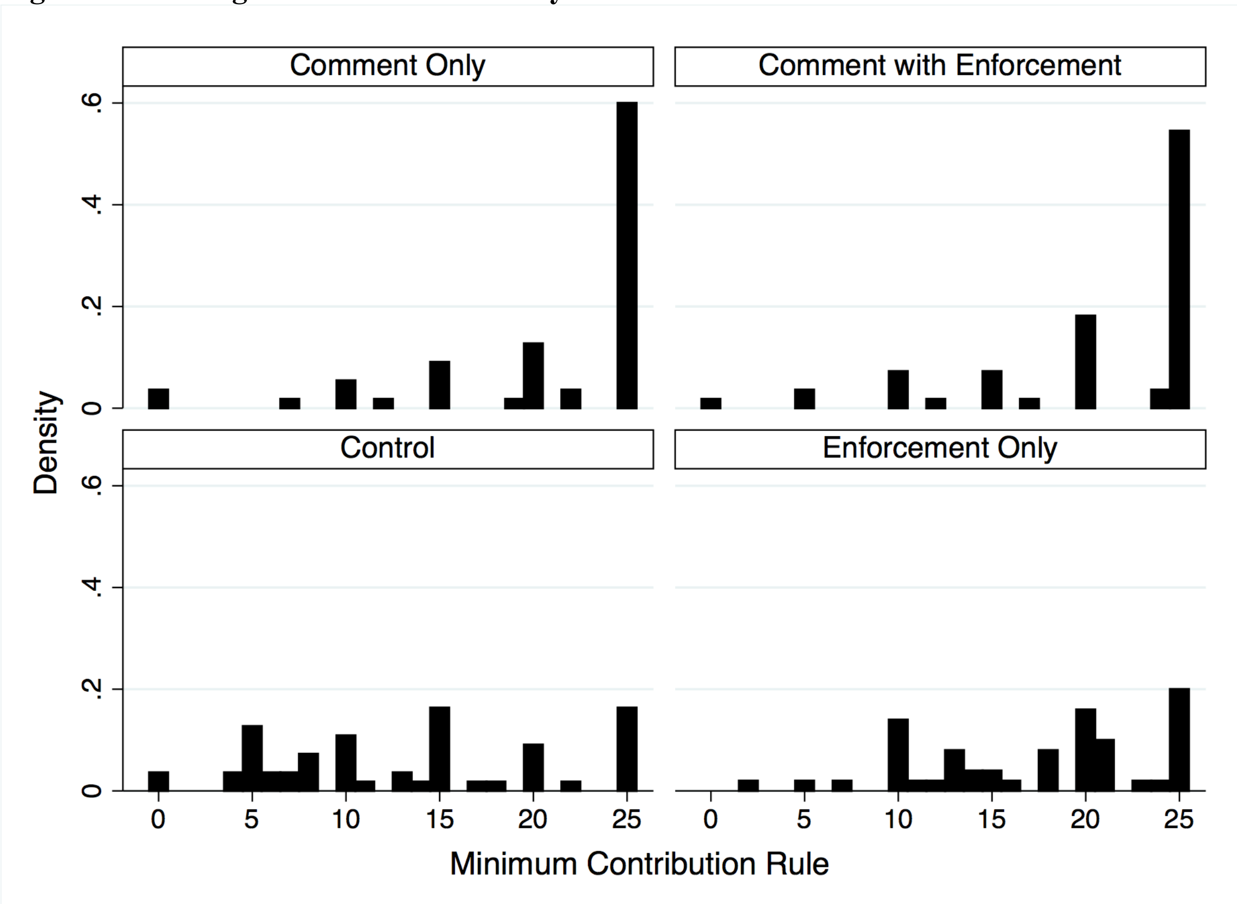


Table 1.5 presents the APEs from Tobit regressions of the MCR chosen by the rule-maker on the treatment indicators and other controls.[10] The associated Tobit coefficient estimates are presented in Appendix Table 1A.5. Across all model specifications, we find strong evidence that exposure to comments increases the level of the MCR chosen by the rule-maker supporting **H4**. These increases move the rule closer to the Pareto optimal contribution rule of 25 points and away from the standard free-riding equilibrium of zero. Focusing on the estimates in model (1), we find that, on average, the comment only treatment increases the chosen MCR by 11.50 points, or 46% of the endowment. Similarly, the comment with enforcement treatment raises the chosen MCR by 10.67

[10] The OLS results are consistent with the Tobit results presented here.

points on average. Here, the enforcement only effect is significantly smaller in magnitude, raising the MCR by only 4.39 points. Additionally, we find that while all treatments result in significantly higher average MCRs than in the control group, the comment only treatment and the comment with enforcement treatment are statistically indistinguishable from one another. This suggests that the effects of comment are similar across institutional structures and outweigh any independent effect of the enforcement regime.

Model (2) investigates the effects of the game order on the choice of an MCR. Here, we see strong evidence of our earlier argument that as the experiment progressed there was some evidence of learning on the part of individuals tasked with choosing the contribution rule. Relative to the first two games, an individual serving in the role of rule-maker chose a rule approximately 5 points higher on average during games 3 through 5. This result helps to explain the significant game effects observed in the noncompliance regressions, where individuals were less likely to comply with rules chosen in later rounds of the game. Because the compliance burden was higher, individuals were less likely to commit the full amount requested to the public good.

We also perform a robustness check to see if previous player experiences in the experiment may be influencing their decisions when they assume the role of rule-maker. Models (4) and (5) present APEs after accounting for the average subject contribution to the public good in the previous game (i.e., when the current rule-maker was a player) as well as the previous MCR.[11] While the impacts of treatment remain unchanged, we observe no significant effect of either player average contributions or the MCR chosen in the last game.

[11] To further test the robustness of these results, we used several measures of lagged experience in the experiment. These measures included previous game earnings, the average of all previous game earnings, and the average of all previous contributions (both player and group). None were statistically significant.

Table 1.5: APEs of treatment on the minimum contribution rule (Tobit estimates)

	(1)	(2)	(3)	(4)
Comment Only (C)	11.504***	11.495***	10.994***	11.655***
	(2.069)	(1.983)	(2.339)	(2.202)
Comment with Enforcement (CE)	10.674***	10.738***	9.566***	10.149***
	(1.983)	(1.884)	(2.535)	(2.407)
Enforcement Only (E)	4.385***	4.350***	3.525*	3.744**
	(1.623)	(1.563)	(1.927)	(1.837)
Game 2		0.569		
		(1.850)		
Game 3		5.141**		4.420**
		(2.191)		(1.978)
Game 4		5.396**		4.960**
		(2.117)		(1.949)
Game 5		5.891**		5.414**
		(2.329)		(2.240)
Previous Game Average Player Contribution			0.141	0.16
			(0.168)	(0.167)
Previous Game Contribution Rule			0.024	-0.065
			(0.122)	(0.124)
Observations	215	215	172	172
Hypotheses		**p-value**		
H_0: C = CE	0.731	0.747	0.633	0.606
H_0: C = E	0.001	0.000	0.003	0.001
H_0: CE = E	0.001	0.001	0.008	0.004

Note: Dependent variable is equal to the level of the MCR chosen by the policymaker (between zero and 25 points). Standard errors in parentheses and clustered at the group level. Significance denoted by *** p<0.01, ** p<0.05, * p<0.1 respectively.

1.5.4 Direct and indirect effect of comments

The strong effects of comments on the rule choice of third-party actors in the VCM setting suggest an alternate pathway through which comments may drive higher contributions to the public good. Specifically, comments may have both direct and indirect effects on contribution levels. Through the direct effect, subjects may change their behavior as a result of the invitation to comment. The indirect effect operates through the MCR. In the previous sub-section, we saw that exposure to comments raises the level of the MCR chosen by a rule-maker (Table 1.5). This higher MCR then has a significant effect on individual contribution levels (Table 1.2).

We decompose the total effect of comments on contribution levels into its direct and indirect components using mediation analysis in Figure 1.5. Mediation analysis is used to investigate alternative causal pathways between treatment and outcome variables by examining the role of intermediate variables (Baron and Kenny 1986; Hayes 2018; Imai et al. 2010). We use the “Medeff” package in Stata to conduct the mediation analysis (Hicks and Tingley 2011). Two caveats should be mentioned. First, the mediation analysis is only available for a single binary treatment variable. We therefore compare the comment only treatment to the control group.[12] Second, there is no mediation analysis estimator for a Tobit model, so we rely on OLS estimation in our mediation analysis.

Overall, Figure 1.5 illustrates that an invitation to submit comments to the rule-maker is found to increase individual contributions to the public good by a total of 3.75 points on average, holding other factors constant.[13] The direct effect of comments on individual behavior accounts for 1.23 points, or 33% of the total effect. The majority of the effect is indirect (operating through the MCR) and raises contributions by 2.52 points or 67% of the total effect.

These results suggest that comments may be an effective means to motivate increased contribution to a public good when stakeholders have the opportunity to influence rule-makers. This finding is significant in light of previous research indicating that bureaucratic rule-makers

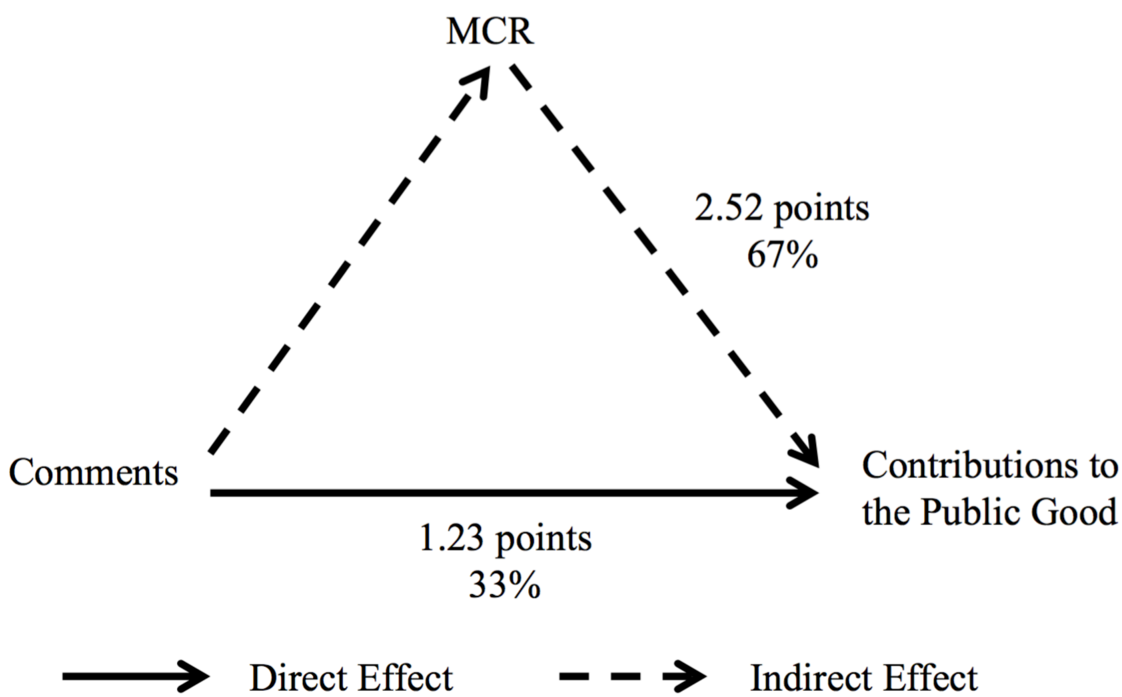
---

[12] As a robustness check we also compared pooled comment treatments (i.e., comment only and comment with enforcement) to pooled no comment treatments (i.e., enforcement only and control). The results of the mediation analysis were the same.

[13] This measure of total treatment effect is similar to our OLS estimate of the effect of the comment only treatment on individual contributions.

and program administrators often respond to comments by modifying the language and rules contained in final policy outcomes (Cropper et al. 1992; Crow et al. 2016).

Figure 1.5: Direct and indirect effects of comments



1.6 Discussion

Governance of public goods and common property resources often incorporate participation elements in an effort to increase cooperation and coordination in overcoming the social dilemma (Ostrom 1990). Participation – through creating increased levels of legitimacy – is then expected to raise individual contributions to public goods and increase individual rates of compliance with regulatory policies. Utilizing a novel lab experiment, this article leverages a public goods contribution mechanism to test the effects of participation – in the form of anonymous comments – on individual contributions and compliance, with and without an exogenous enforcement mechanism. By including comment treatments with and without enforcement, we are able to investigate potential interactions between these two elements of policy formulation and implementation. Unlike previous experiments where the contribution rule is determined by voting (e.g., majority rule), the contribution rule here is determined by a third party acting as the rule-maker. This type of mechanism better reflects the formation of regulatory policy in practice - e.g., individuals who comment are attempting to influence the final outcome of the regulation but have no guarantee that their comments will be taken into account. Additionally, this element of the design provided leverage to understand how the receipt of comments influences the rule-maker's choice of contribution rule.

Our results show significant evidence that the implemented type of comment participation can increase contributions to a public good while reducing both the probability that an individual will engage in noncompliant behavior and the extent of noncompliant behavior, but this impact is conditional on the presence of an exogenous enforcement mechanism. Specifically, we find that, relative to treatments with no comments or enforcement, comments alone were successful in raising contributions during early rounds of each game to levels similar to enforcement only treatments. However, after several rounds, contributions decayed rapidly dropping to levels similar to the no comments or enforcement treatment. This suggests that comments, in isolation, may not be sufficient to sustain increased contributions. Combining comments with enforcement, however, does lead to higher and sustained contributions, even more so than when there is enforcement only. It appears that there is an important interaction effect where comments significantly increase contributions (above even enforcement only) and enforcement leads to these higher contributions being sustained over time. Kroll et al. (2007) and Messer et al. (2007) find similar results for the effectiveness of nonbinding voting suggesting that complementary sanctioning or coordinating mechanisms may help to achieve gains on public good outcomes.

We find similar results regarding rates of compliance. Here, comments alone do not motivate increased rates of compliance, but when individuals are exposed to comments with enforcement, we observe significant reductions beyond enforcement alone in the probability that an individual engages in noncompliance as well as the extent of noncompliance. Looking at the behavior of the third-party rule-maker, we find that receiving comments significantly increases the rule-maker's choice of an MCR, raising it closer to the Pareto optimum. This effect holds regardless of whether or not an enforcement regime is imposed during the play of game. A higher MCR, in turn, increases individual contributions to the public good. Overall, comments have a significant but small direct effect on individual contributions to the public good, and a larger indirect effect operating through a higher contribution rule.

This research has several policy implications. First, it suggests that participation mechanisms might be a viable means for increasing contributions to public goods while reducing noncompliance as long as enforcement mechanisms are also in place. Having this complementary institution available allows individuals to sustain coordination at higher levels of public good provision than might be achieved through comment alone. Because of this complementarity, the additional gains from participation may only be realized in communities with high levels of institutional capacity for monitoring and enforcement. Comments are not a substitute for, but rather a complement to, compliance campaigns or schemes. Second, comments may be a vehicle for cost savings on monitoring and enforcement budgets. Given that most comments are handled through streamlined online platforms, the costs of incorporating additional participation opportunities is likely to be low. Our findings suggest that certain enforcement costs (e.g., litigation) might fall if there are lower numbers of individuals engaging in noncompliance as a result of the comment mechanism. However, if the threat of sanctioning for non-compliers must be relatively high to encourage compliance, monitoring costs may not decrease. More research is needed to fully understand the costs of implementing a comment mechanism vis-à-vis current monitoring and enforcement regimes. Third, we find fairly large effects of comments on the choice of rule by our rule-makers that appear to be independent of the presence of an enforcement mechanism. Increased democratization of policymaking processes through comment mechanisms, even if they do not require tying the hands of rule-makers through a commitment to majority rule voting, may be able to help rule-makers achieve higher levels of public goods provision than those chosen without participant input. Finally, our work suggests that the participatory nature of rulemaking institutions can have significant effects on individual propensity to comply with voluntary standards (Ostrom 1990; Walker et al. 2000).

Future research promises several worthwhile extensions. One would be to implement comments under different institutional structures. A key specification would be under an endogenous enforcement regime that would have a game theoretic solution where individuals can sanction one another (Fehr and Gächter 2000; Kamei et al. 2015; Putterman et al. 2011). Examples might be that only the lowest contributor to the public good would be sanctioned or that sanctioning would only apply to individuals who contribute less than some proportion of the minimum contribution rule. Another important area of research is to understand the effectiveness of the comment mechanism in the presence of individual heterogeneity over preferences for the public good as well as endowments. Heterogeneity has been an important driver of institutional effectiveness in the public goods setting (de Oliveira et al. 2015; Kesternich et al. 2018). Our implementation necessarily places all individuals on the same side of the social dilemma where higher levels of provision are socially optimal. Having groups of people who might gain differential benefits or bear disproportionate costs may change the effect of the comment mechanism.

APPENDICES

APPENDIX 1A: Robustness checks

Table 1A.1: Summary statistics for key outcomes of interest

Panel A: Pooled averages by treatment for key outcomes

Treatments	Contribution	Binary Noncompliance	Relative Noncompliance	Minimum Contribution Rule	Payoff	Observations
Control	10.64 (8.537)	0.445 (0.497)	0.317 (0.402)	13.25 (7.355)	31.38 (6.761)	1100
Comment Only	14.37 (9.912)	0.498 (0.500)	0.335 (0.409)	20.95 (6.389)	33.62 (8.048)	1100
Enforcement Only	17.02 (7.483)	0.157 (0.364)	0.114 (0.293)	17.34 (6.068)	33.26 (7.558)	1000
Comment with Enforcement	20.27 (7.465)	0.086 (0.280)	0.068 (0.244)	20.67 (6.299)	36.11 (6.394)	1100

Note: Standard errors in parentheses. Individual data was pooled across all games and rounds.

Panel B: Hypothesis testing between pooled means

Null Hypothesis	Contribution	Binary Noncompliance	Relative Noncompliance	Minimum Contribution Rule	Payoff
Control = Comment	0.000	0.012	0.297	0.000	0.000
Control = Enforcement	0.000	0.000	0.000	0.000	0.000
Control = Comment with Enforcement	0.000	0.000	0.000	0.000	0.000
Comment = Enforcement	0.000	0.000	0.000	0.000	0.295
Comment = Comment with Enforcement	0.000	0.000	0.000	0.314	0.000
Enforcement = Comment with Enforcement	0.000	0.000	0.000	0.000	0.000

Note: P-values reported from two-sided test of differences in means

Table 1A.2: Tobit coefficient estimates for contributions to the group account

	(1)	(2)	(3)	(4)	(5)
Minimum Contribution Rule	0.747***	0.746***	0.745***	0.789***	0.791***
	(0.078)	(0.078)	(0.078)	(0.079)	(0.080)
Comment Only (C)	-0.619	-0.609	3.122**	-0.948	5.028*
	(1.774)	(1.774)	(1.395)	(1.754)	(3.026)
Comment with Enforcement (CE)	10.689***	10.654***	8.892***	10.249***	3.368
	(1.684)	(1.679)	(1.508)	(1.679)	(2.825)
Enforcement Only (E)	5.203***	5.193***	2.614**	5.033***	2.653
	(1.370)	(1.367)	(1.145)	(1.356)	(2.364)
Round 2		-0.603	-0.643	-0.596	-0.639
		(0.375)	(0.542)	(0.374)	(0.534)
Round 3		-1.622***	-1.921**	-1.609***	-1.897**
		(0.491)	(0.816)	(0.490)	(0.805)
Round 4		-2.213***	-2.343***	-2.204***	-2.319***
		(0.558)	(0.902)	(0.557)	(0.891)
Round 5		-5.176***	-5.058***	-5.163***	-5.031***
		(0.693)	(1.017)	(0.692)	(1.008)
C x Round 2			-1.988**		-1.895**
			(0.940)		(0.948)
C x Round 3			-3.565***		-3.491**
			(1.376)		(1.380)
C x Round 4			-5.204***		-5.191***
			(1.543)		(1.543)
C x Round 5			-7.769***		-7.827***
			(1.751)		(1.757)
CE x Round 2			1.783*		1.852*
			(0.969)		(0.992)
CE x Round 3			1.875*		1.817
			(1.138)		(1.148)
CE x Round 4			2.547**		2.504**
			(1.178)		(1.179)
CE x Round 5			2.270		2.175
			(1.480)		(1.486)
E x Round 2			0.537		0.552
			(0.966)		(0.962)
E x Round 3			3.159***		3.133***
			(1.162)		(1.153)
E x Round 4			3.619***		3.595***
			(1.285)		(1.276)
E x Round 5			5.542***		5.535***
			(1.326)		(1.320)
Game 2				0.328	0.439
				(1.516)	(3.287)
Game 3				-2.307	-1.491
				(1.688)	(3.504)
Game 4				-1.593	-3.902
				(1.750)	(3.613)
Game 5				-2.482	-3.994
				(1.696)	(3.436)

Table 1A.2 (cont'd)

	(1)	(2)	(3)	(4)	(5)
C x Game 2					0.572
					(4.533)
C x Game 3					-7.216
					(5.160)
C x Game 4					-2.368
					(5.355)
C x Game 5					-2.638
					(4.966)
CE x Game 2					0.917
					(4.008)
CE x Game 3					6.503
					(4.668)
CE x Game 4					10.139**
					(4.621)
CE x Game 5					9.569**
					(4.855)
E x Game 2					-1.804
					(3.739)
E x Game 3					-1.948
					(3.928)
E x Game 4					2.584
					(4.119)
E x Game 5					0.073
					(3.791)
Constant	0.589	2.516*	2.605**	3.142**	3.762*
	(1.458)	(1.419)	(1.232)	(1.573)	(2.264)
Observations	4,300	4,300	4,300	4,300	4,300

Note: Dependent variable is the number of points contributed to the group account from zero to 25. Standard errors in parentheses and clustered at the group level. Significance denoted by *** p<0.01, ** p<0.05, * p<0.1 respectively

Table 1A.3: Probit coefficient estimates for binary noncompliance

	(1)	(2)	(3)	(4)	(5)
Comment Only (C)	0.135	0.137	-0.050	0.139	-0.254
	(0.150)	(0.152)	(0.141)	(0.148)	(0.324)
Comment with Enforcement (CE)	-1.230***	-1.244***	-1.014***	-1.256***	-0.671**
	(0.147)	(0.149)	(0.172)	(0.152)	(0.315)
Enforcement Only (E)	-0.867***	-0.871***	-0.544***	-0.885***	-0.639**
	(0.140)	(0.142)	(0.161)	(0.141)	(0.315)
Round 2		0.107**	0.097	0.108**	0.099
		(0.050)	(0.060)	(0.052)	(0.062)
Round 3		0.241***	0.319***	0.245***	0.326***
		(0.054)	(0.068)	(0.056)	(0.071)
Round 4		0.317***	0.410***	0.321***	0.420***
		(0.054)	(0.072)	(0.055)	(0.073)
Round 5		0.476***	0.513***	0.486***	0.525***
		(0.061)	(0.079)	(0.063)	(0.081)
C x Round 2			0.168*		0.176*
			(0.094)		(0.099)
C x Round 3			0.221**		0.241**
			(0.107)		(0.113)
C x Round 4			0.176		0.195*
			(0.113)		(0.117)
C x Round 5			0.370***		0.405***
			(0.132)		(0.137)
CE x Round 2			-0.236		-0.235
			(0.173)		(0.173)
CE x Round 3			-0.288*		-0.288*
			(0.151)		(0.153)
CE x Round 4			-0.410***		-0.411***
			(0.152)		(0.154)
CE x Round 5			-0.207		-0.208
			(0.163)		(0.165)
E x Round 2			-0.059		-0.067
			(0.137)		(0.142)
E x Round 3			-0.540***		-0.555***
			(0.149)		(0.151)
E x Round 4			-0.430***		-0.449***
			(0.138)		(0.141)
E x Round 5			-0.595***		-0.609***
			(0.154)		(0.159)
Game 2				0.010	-0.024
				(0.159)	(0.334)
Game 3				0.386**	0.322
				(0.176)	(0.358)
Game 4				0.350**	0.276
				(0.164)	(0.371)
Game 5				0.447***	0.544
				(0.167)	(0.369)
C x Game 2					0.054
					(0.424)
C x Game 3					0.327
					(0.499)

Table 1A.3 (cont'd)

	(1)	(2)	(3)	(4)	(5)
C x Game 4					0.433
					(0.470)
C x Game 5					0.129
					(0.481)
CE x Game 2					-0.249
					(0.442)
CE x Game 3					-0.273
					(0.456)
CE x Game 4					-0.386
					(0.445)
CE x Game 5					-0.859*
					(0.452)
E x Game 2					0.327
					(0.410)
E x Game 3					0.063
					(0.448)
E x Game 4					0.048
					(0.429)
E x Game 5					0.026
					(0.438)
Constant	-0.139	-0.370***	-0.410***	-0.615***	-0.644***
	(0.115)	(0.117)	(0.110)	(0.156)	(0.244)
Observations	4,300	4,300	4,300	4,300	4,300

Note: Dependent variable is equal to one if the subject does not contribute at least the MCR to the group account and is zero otherwise. Standard errors in parentheses and clustered at the group level. Significance denoted by *** $p<0.01$, ** $p<0.05$, * $p<0.1$ respectively.

Table 1A.4:Fractional response coefficient estimates for relative noncompliance

	(1)	(2)	(3)	(4)	(5)
Comment Only (C)	0.050	0.049	-0.126	0.047	-0.246
	(0.133)	(0.136)	(0.123)	(0.133)	(0.304)
Comment with Enforcement (CE)	-1.019***	-1.036***	-0.842***	-1.043***	-0.491*
	(0.141)	(0.144)	(0.161)	(0.146)	(0.291)
Enforcement Only (E)	-0.732***	-0.734***	-0.462***	-0.745***	-0.555*
	(0.126)	(0.128)	(0.137)	(0.126)	(0.294)
Round 2		0.103**	0.080	0.104**	0.079
		(0.043)	(0.057)	(0.044)	(0.058)
Round 3		0.216***	0.295***	0.219***	0.296***
		(0.048)	(0.067)	(0.049)	(0.069)
Round 4		0.323***	0.362***	0.327***	0.367***
		(0.048)	(0.067)	(0.049)	(0.068)
Round 5		0.543***	0.566***	0.551***	0.572***
		(0.055)	(0.075)	(0.055)	(0.076)
C x Round 2			0.122		0.124
			(0.075)		(0.077)
C x Round 3			0.133		0.146
			(0.095)		(0.098)
C x Round 4			0.233**		0.250**
			(0.101)		(0.103)
C x Round 5			0.339***		0.365***
			(0.110)		(0.113)
CE x Round 2			-0.199		-0.203
			(0.158)		(0.156)
CE x Round 3			-0.275**		-0.272**
			(0.133)		(0.135)
CE x Round 4			-0.320**		-0.322**
			(0.140)		(0.142)
CE x Round 5			-0.162		-0.164
			(0.151)		(0.153)
E x Round 2			0.046		0.045
			(0.126)		(0.131)
E x Round 3			-0.429***		-0.437***
			(0.144)		(0.145)
E x Round 4			-0.356***		-0.368***
			(0.128)		(0.131)
E x Round 5			-0.580***		-0.585***
			(0.155)		(0.158)
Game 2				0.063	0.172
				(0.152)	(0.287)
Game 3				0.378**	0.299
				(0.158)	(0.297)
Game 4				0.347**	0.389
				(0.163)	(0.320)
Game 5				0.384**	0.385
				(0.153)	(0.293)
C x Game 2					-0.155
					(0.394)
C x Game 3					0.328
					(0.427)
C x Game 4					0.162
					(0.449)
C x Game 5					0.149
					(0.422)

Table 1A.4 (cont'd)

	(1)	(2)	(3)	(4)	(5)
CE x Game 2					-0.445
					(0.434)
CE x Game 3					-0.211
					(0.435)
CE x Game 4					-0.582
					(0.432)
CE x Game 5					-0.554
					(0.401)
E x Game 2					0.083
					(0.390)
E x Game 3					0.068
					(0.392)
E x Game 4					0.020
					(0.414)
E x Game 5					0.224
					(0.385)
Constant	-0.475***	-0.720***	-0.745***	-0.964***	-1.002***
	(0.097)	(0.097)	(0.088)	(0.137)	(0.206)
Observations	4,300	4,300	4,300	4,300	4,300

Note: Dependent variable is equal to the proportion of the MCR not accounted for by the individual contribution to the group account. Standard errors in parentheses and clustered at the group level. Significance denoted by *** p<0.01, ** p<0.05, * p<0.1 respectively.

Table 1A.5: Tobit coefficient estimates for the effect of treatment on the minimum contribution rule

	(1)	(2)	(3)	(4)
Comment Only (C)	11.856***	11.843***	11.221***	11.905***
	(2.121)	(2.027)	(2.379)	(2.238)
Comment with Enforcement (CE)	11.021***	11.082***	9.788***	10.394***
	(2.046)	(1.939)	(2.598)	(2.469)
Enforcement Only (E)	4.630***	4.592***	3.669*	3.907**
	(1.720)	(1.654)	(2.010)	(1.920)
Game 2		0.600		
		(1.952)		
Game 3		5.319**		4.565**
		(2.260)		(2.026)
Game 4		5.579**		5.115**
		(2.192)		(2.008)
Game 5		6.082**		5.575**
		(2.391)		(2.287)
Previous Game Average Player Contribution			0.144	0.164
			(0.173)	(0.171)
Previous Game Contribution Rule			0.024	-0.067
			(0.125)	(0.127)
Observations	215	215	172	172

Note: Dependent variable is equal to the level of the MCR chosen by the policymaker (between zero and 25 points). Standard errors in parentheses and clustered at the group level. Significance denoted by *** p<0.01, ** p<0.05, * p<0.1 respectively.

APPENDIX 1B: Experimental instructions

At the beginning of each experimental session, subjects participated in a practice game. The practice game was the same for each experimental treatment. Please find the practice game instructions below.

Practice Game Instructions

This is an experiment about individual and group decision-making. The amount of money you earn depends on the decisions that you and the other participants make.

You will never be asked to reveal your identity to anyone during the course of the experiment. Your name will never be associated with any of your decisions. In order to keep your decisions private do not reveal your choices to any other participant. Please do not talk to one another for the duration of the experiment. If you have any questions, please raise your hand.

1. There will be 5 rounds in the practice game.

2. You will be randomly assigned to groups of 5 players

3. At the beginning of each round, you will receive 25 points.

4. In each round, you will choose how many points you want to place in a group account. Any points not placed in the group account will be kept in your private account. The total

number of points placed in your private and group account will add up to 25 points.

5. You get to keep the full value of any points allocated to your private account. For each point allocated to the group account, you will earn 0.4 points and each other person in your group will earn 0.4 points. This means that each point placed in the group account returns 2.0 points to the group as a whole.

6. After all players have made an allocation decision, you will be informed of your total earnings for that round. You will also be informed of the anonymous decisions of your other group members.

7. During the experiment, you are not permitted to speak or communicate with other participants. If you have a question while the practice game is going on, please raise your hand and one of us will come to your seat to answer it. Practice game points do not count towards your final payment. At this time, do you have any questions about the instructions or procedures?

Following the practice game, research subjects participated in one of four experimental treatments. The treatments were:

1. Control (no comment and no enforcement)
2. Comment only
3. Enforcement only

4. Comment with enforcement

Treatment 1 Instructions: Control

1. This experiment consists of five individual games. At the beginning of each game, you will be randomly assigned to a group of five (you plus four other people). Each game consists of five rounds. You will be matched with the same people for all five rounds of a game. Your earnings in each round will depend upon the decisions that you make and the decisions that the other people in your group make.

2. At the beginning of each game, you will be randomly assigned a role. The first role is referred to as **Player**. There will be 4 **Players** in each group. The second role is referred to as the **Policymaker**. Each group will have a single **Policymaker**. Your role will remain the same for all five rounds of a game. Each participant will have the opportunity to play the role of **Policymaker** at least once today.

3. Just as in the practice game, **Players** will be given 25 points and asked to divide those points between a group account and an individual account. You will choose the number of points you wish to put in the group account. Any points not placed in the group account will be placed in your private account. Thus, the total number of points placed in your private and group account will add up to 25 points.

4. You get to keep any points allocated to your private account. For each point allocated to

the group account, you will earn .4 points and each other **Player** in your group will earn .4 points. This means that each point placed in the group account returns 1.6 points to the group as a whole at the end of a round. Note, this is different from the practice game because there are now 4 **Players** instead of 5. After all **Players** have made a contribution decision, you will be informed of your total earnings for that round. You will also be informed of the decisions of each member of the group while maintaining anonymity.

5. It is the job of the **Policymaker** to choose a minimum allocation rule. The minimum allocation rule is the minimum number of points that the **Policymaker** believes each **Player** should put in the group account during each round. The **Policymaker** can choose any amount between 0 and 25 points for the minimum allocation rule. Note that the **Policymaker** will be the same for all five rounds of each game. Additionally, the minimum allocation rule chosen at the beginning of the game will be the same for all five rounds. For completing the task of choosing a minimum allocation rule, the **Policymaker** will be paid a salary of 25 points per round.

6. We will record your point earnings during every round. At the end of the experiment we will add up your point earnings and convert them to U.S. dollars at a rate of $3.00 per 100 points. We will pay you this amount privately and in cash. Your earnings are your own business and you do not have to discuss them with anyone.

7. During the experiment, you are not permitted to speak or communicate with the other participants. If you have a question while the experiment is going on, please raise your

hand and one of us will come to your seat to answer it. At this time, do you have any questions about the instructions or procedures?

Treatment 2 Instructions: Comment only

1. This experiment consists of five individual games. At the beginning of each game, you will be randomly assigned to a group of five (you plus four other people). Each game consists of five rounds. You will be matched with the same people for all five rounds of a game. Your earnings in each round will depend upon the decisions that you make and the decisions that the other people in your group make.

2. At the beginning of each game, you will be randomly assigned a role. The first role is referred to as **Player**. There will be 4 **Players** in each group. The second role is referred to as the **Policymaker**. Each group will have a single **Policymaker**. Your role will remain the same for all five rounds of a game. Each participant will have the opportunity to play the role of **Policymaker** at least once today.

3. Just as in the practice game, **Players** will be given 25 points, and asked to divide those points between a group account and an individual account. You will choose the number of points you wish to put in the group account. Any points not placed in the group account will be placed in your private account. Thus, the total number of points placed in your private and group account will add up to 25 points.

4. You get to keep any points allocated to your private account. For each point allocated to the group account, you will earn .4 points and each other **Player** in your group will earn .4

points. This means that each point placed in the group account returns 1.6 points to the group as a whole at the end of a round. Note, this is different from the practice game because there are now 4 **Players** instead of 5. After all **Players** have made a contribution decision, you will be informed of your total earnings for that round. You will also be informed of the decisions of each member of the group while maintaining anonymity.

5. At the beginning of each game, all **Players** will be asked to provide comments about their preferences for a minimum allocation rule. The minimum allocation rule is the minimum number of points that the **Policymaker** believes each **Player** should put in the group account during each round. The minimum allocation rule can be any amount from 0 to 25 points. You will have time to enter any comments you wish. Note that your comments may or may not influence the **Policymaker's** choice of a minimum allocation rule.

6. After all **Players** have submitted their comments about the minimum allocation rule, these comments will be provided to the **Policymaker**. It is the job of the **Policymaker** to read the comments and decide what the minimum allocation rule should be. The **Policymaker** can choose any amount between 0 and 25 points for the minimum allocation rule. Note that the **Policymaker** will be the same for all five rounds of each game. Additionally, the minimum allocation rule chosen at the beginning of the game will be the same for all five rounds. For completing the task of reading the comments and choosing a minimum allocation rule, the **Policymaker** will be paid a salary of 25 points per round.

7. We will record your point earnings during every round. At the end of the experiment we will add up your point earnings and convert them to U.S. dollars at a rate of $3.00 per 100

points. We will pay you this amount privately and in cash. Your earnings are your own business and you do not have to discuss them with anyone.

8. During the experiment, you are not permitted to speak or communicate with the other participants. If you have a question while the experiment is going on, please raise your hand and one of us will come to your seat to answer it. At this time, do you have any questions about the instructions or procedures?

<u>Treatment 3 Instructions: Enforcement only</u>

1. This experiment consists of five individual games. At the beginning of each game, you will be randomly assigned to a group of five (you plus four other people). Each game consists of five rounds. You will be matched with the same people for all five rounds of a game. Your earnings in each round will depend upon the decisions that you make and the decisions that the other people in your group make.

2. At the beginning of each game, you will be randomly assigned a role. The first role is referred to as **Player**. There will be 4 **Players** in each group. The second role is referred to as the **Policymaker**. Each group will have a single **Policymaker**. Your role will remain the same for all five rounds of a game. Each participant will have the opportunity to play the role of **Policymaker** at least once today.

3. Just as in the practice game, **Players** will be given 25 points and asked to divide those points between a group account and an individual account. You will choose the number of

points you wish to put in the group account. Any points not placed in the group account will be placed in your private account. Thus, the total number of points placed in your private and group account will add up to 25 points.

4. You get to keep any points allocated to your private account. For each point allocated to the group account, you will earn .4 points and each other **Player** in your group will earn .4 points. This means that each point placed in the group account returns 1.6 points to the group as a whole at the end of a round. Note, this is different from the practice game because there are now 4 **Players** instead of 5. After all **Players** have made a contribution decision, you will be informed of your total earnings for that round. You will also be informed of the decisions of each member of the group while maintaining anonymity.

5. It is the job of the **Policymaker** to choose a minimum allocation rule. The minimum allocation rule is the minimum number of points that the **Policymaker** believes each **Player** should put in the group account during each round. The **Policymaker** can choose any amount between 0 and 25 points for the minimum allocation rule. Note that the **Policymaker** will be the same for all five rounds of each game. Additionally, the minimum allocation rule chosen at the beginning of the game will be the same for all five rounds. For completing the task of choosing a minimum allocation rule, the **Policymaker** will be paid a salary of 25 points per round.

6. In this experiment, **Players** who do not contribute at least the minimum allocation rule to the group account may be fined. The probability of being caught is 50% and the fine is 25

points. This means that if you contribute less than the minimum allocation rule in any round, you have a 50% chance of having 25 points subtracted from your overall earnings.

7. We will record your point earnings during every round. At the end of the experiment we will add up your point earnings and convert them to U.S. dollars at a rate of $3.00 per 100 points. We will pay you this amount privately and in cash. Your earnings are your own business and you do not have to discuss them with anyone.

8. During the experiment, you are not permitted to speak or communicate with the other participants. If you have a question while the experiment is going on, please raise your hand and one of us will come to your seat to answer it. At this time, do you have any questions about the instructions or procedures?

<u>Treatment 4 Instructions: Comment with enforcement</u>

1. This experiment consists of five individual games. At the beginning of each game, you will be randomly assigned to a group of five (you plus four other people). Each game consists of five rounds. You will be matched with the same people for all five rounds of a game. Your earnings in each round will depend upon the decisions that you make and the decisions that the other people in your group make.

2. At the beginning of each game, you will be randomly assigned a role. The first role is referred to as **Player**. There will be 4 **Players** in each group. The second role is referred to as the **Policymaker**. Each group will have a single **Policymaker**. Your role will remain

the same for all five rounds of a game. Each participant will have the opportunity to play the role of **Policymaker** at least once today.

3. Just as in the practice game, **Players** will be given 25 points, and asked to divide those points between a group account and an individual account. You will choose the number of points you wish to put in the group account. Any points not placed in the group account will be placed in your private account. Thus, the total number of points placed in your private and group account will add up to 25 points.

4. You get to keep any points allocated to your private account. For each point allocated to the group account, you will earn .4 points and each other **Player** in your group will earn .4 points. This means that each point placed in the group account returns 1.6 points to the group as a whole at the end of a round. Note, this is different from the practice game because there are now 4 **Players** instead of 5. After all **Players** have made a contribution decision, you will be informed of your total earnings for that round. You will also be informed of the decisions of each member of the group while maintaining anonymity.

5. At the beginning of each game, all **Players** will be asked to provide comments about their preferences for a minimum allocation rule. The minimum allocation rule is the minimum number of points that the **Policymaker** believes each **Player** should put in the group account during each round. The minimum allocation rule can be any amount from 0 to 25 points. You will have time to enter any comments you wish. Note that your comments may or may not influence the **Policymaker's** choice of a minimum allocation rule.

6. After all **Players** have submitted their comments about the minimum allocation rule, these comments will be provided to the **Policymaker**. It is the job of the **Policymaker** to read the comments and decide what the minimum allocation rule should be. The **Policymaker** can choose any amount between 0 and 25 points for the minimum allocation rule. Note that the **Policymaker** will be the same for all five rounds of each game. Additionally, the minimum allocation rule chosen at the beginning of the game will be the same for all five rounds. For completing the task of reading the comments and choosing a minimum allocation rule, the **Policymaker** will be paid a salary of 25 points per round.

7. In this experiment, **Players** who do not contribute at least the minimum allocation rule to the group account may be fined. The probability of being caught is 50% and the fine is 25 points. This means that if you contribute less than the minimum allocation rule in any round, you have a 50% chance of having 25 points subtracted from your overall earnings.

8. We will record your point earnings during every round. At the end of the experiment we will add up your point earnings and convert them to U.S. dollars at a rate of $3.00 per 100 points. We will pay you this amount privately and in cash. Your earnings are your own business and you do not have to discuss them with anyone.

9. During the experiment, you are not permitted to speak or communicate with the other participants. If you have a question while the experiment is going on, please raise your

hand and one of us will come to your seat to answer it. At this time, do you have any questions about the instructions or procedures?

REFERENCES

REFERENCES

Anderson, C. M., & Putterman, L. (2006). Do non-strategic sanctions obey the law of demand? The demand for punishment in the voluntary contribution mechanism. *Games and Economic Behavior*, *54*(1), 1–24. doi:10.1016/j.geb.2004.08.007

Anderson, L. R., & Stafford, S. L. (2003). Punishment in a Regulatory Setting: Experimental Evidence from the VCM. *Journal of Regulatory Economics*, *24*(1), 91–110.

Andreoni, J. (1995). Warm-Glow versus Cold-Prickle: The Effects of Positive and Negative Framing on Cooperation in Experiments. *The Quarterly Journal of Economics*, *110*(1), 1–21. doi:10.2307/2118508

Andreoni, J. (2006). Philanthropy. In *Handbook of the Economics of Giving, Altruism, and Reciprocity: Applications* (Vol. 2). North-Holland. http://econweb.ucsd.edu/~jandreon/WorkingPapers/Philanthropy.pdf. Accessed 28 May 2018

Bardhan, P. (2000). Irrigation and Cooperation: An Empirical Analysis of 48 Irrigation Communities in South India. *Economic Development and Cultural Change*, *48*(4), 847–865. doi:10.1086/452480

Baron, R. M., & Kenny, D. A. (1986). The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182.

Bernhard, H., Fischbacher, U., & Fehr, E. (2006). Parochial altruism in humans. *Nature*, *442*(7105), 912–915. doi:10.1038/nature04981

Black, S.E., & Lynch, L. M. (2001). How to Compete: The Impact of Workplace Practices and Information Technology on Productivity. *The Review of Economics and Statistics*, *83*(3), 434–445.

Bochet, O., Page, T., & Putterman, L. (2006). Communication and punishment in voluntary contribution experiments. *Journal of Economic Behavior & Organization*, *60*(1), 11–26. doi:10.1016/j.jebo.2003.06.006

Bouma, J. A., Joy, K. J., Paranjape, S., & Ansink, E. (2014). The Influence of Legitimacy Perceptions on Cooperation – A Framed Field Experiment. *World Development*, *57*, 127–137. doi:10.1016/j.worlddev.2013.12.007

Bouma, J., & Ansink, E. (2013). The role of legitimacy perceptions in self-restricted resource use: A framed field experiment. *Forest Policy and Economics*, *37*, 84–93. doi:10.1016/j.forpol.2013.01.006

Burlando, R. M., & Guala, F. (2005). Heterogeneous Agents in Public Goods Experiments. *Experimental Economics*, *8*(1), 35–54. doi:10.1007/s10683-005-0436-4

Carpenter, J.P., & Matthews, P. H. (2012). Norm Enforcement: Anger, Indignation, or Reciprocity? *Journal of the European Economic Association*, *10*(3), 555–572.

Charness, G., Cobo-Reyes, R., & Jiménez, N. (2008). An investment game with third-party intervention. *Journal of Economic Behavior & Organization*, *68*(1), 18–28. doi:10.1016/j.jebo.2008.02.006

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics*, *14*(1), 47–83. doi:10.1007/s10683-010-9257-1

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97. doi:10.1016/j.jbef.2015.12.001

Cropper, M. L., Evans, W. N., Berardi, S. J., Ducla-Soares, M. M., & Portney, P. R. (1992). The Determinants of Pesticide Regulation: A Statistical Analysis of EPA Decision Making *Journal of Political Economy*, *100*(1), 175–197.

Croson, R. T. A. (2000). Thinking like a game theorist: factors affecting the frequency of equilibrium play. *Journal of Economic Behavior & Organization*, *41*(3), 299–314. doi:10.1016/S0167-2681(99)00078-5

Croson, R. T. A. (2007). Theories of Commitment, Altruism, and Reciprocity: Evidence from Linear Public Goods Games. *Economic Inquiry*, *45*(2), 199–216. doi:10.1111/j.1465-7295.2006.00006.x

Croson, R. T. A., & Marks, M. B. (2000). Step returns in threshold public goods: A meta- and experimental analysis. *Experimental Economics*, *2*(3), 239–259. doi:10.1007/BF01669198

Crow, D. A., Albright, E. A., & Koebele, E. (2016). Public Information and Regulatory Processes: What the Public Knows and Regulators Decide: Public Information and Regulatory Processes. *Review of Policy Research*, *33*(1), 90–109. doi:10.1111/ropr.12154

Crumpler, H., & Grossman, P. J. (2008). An experimental test of warm glow giving. *Journal of Public Economics*, *92*(5–6), 1011–1021. doi:10.1016/j.jpubeco.2007.12.014

Dal Bó, P., Foster, A., & Putterman, L. (2010). Institutions and Behavior: Experimental Evidence on the Effects of Democracy. *American Economic Review*, *100*(5), 2205–2229. doi:10.1257/aer.100.5.2205

Dannenberg, A., Lange, A., & Sturm, B. (2014). Participation and Commitment in Voluntary Coalitions to Provide Public Goods. *Economica*, *81*(322), 257–275. doi:10.1111/ecca.12073

de Oliveira, A. C. M., Croson, R. T. A., & Eckel, C. (2015). One bad apple? Heterogeneity and information in public good provision. *Experimental Economics*, *18*(1), 116–135. doi:10.1007/s10683-014-9412-1

Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior*, *25*(2), 63–87. doi:10.1016/S1090-5138(04)00005-4

Fehr, E., & Gächter, S. (2000). Cooperation and Punishment in Public Goods Experiments. *American Economic Review*, *90*(4), 980–994. doi:10.1257/aer.90.4.980

Feld, L. P., & Frey, B. S. (2007). Tax Compliance as the Result of a Psychological Tax Contract: The Role of Incentives and Responsive Regulation. *Law & Policy*, *29*(1), 102–120. doi:10.1111/j.1467-9930.2007.00248.x

Feld, L. P., & Tyran, J.-R. (2002). Tax Evasion and Voting: An Experimental Analysis. *Kyklos*, *55*(2), 197–221. doi:10.1111/1467-6435.00183

Fiedler, M., & Haruvy, E. (2017). The effect of third party intervention in the trust game. *Journal of Behavioral and Experimental Economics*, *67*, 65–74. doi:10.1016/j.socec.2016.10.003

Gächter, S., Renner, E., & Sefton, M. (2008). The Long-Run Benefits of Punishment. *Science, New Series*, *322*(5907), 1510.

Gallier, C., Kesternich, M., & Sturm, B. (2017). Voting for Burden Sharing Rules in Public Goods Games. *Environmental and Resource Economics*, *67*(3), 535–557. doi:10.1007/s10640-016-0022-6

Gatiso, T. T., & Vollan, B. (2017). Democracy and cooperation in commons management: experimental evidence of representative and direct democracy from community forests in Ethiopia. *Environment and Development Economics*, *22*(02), 110–132. doi:10.1017/S1355770X16000322

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association*, *1*(1), 114–125. doi:10.1007/s40881-015-0004-4

Gürerk, Ö., Irlenbusch, B., & Rockenbach, B. (2006). The Competitive Advantage of Sanctioning Institutions. *Science, New Series*, *312*(5770), 108–111.

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: a regression-based approach* (Second edition.). New York: Guilford Press.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., et al. (2006). Costly Punishment across Human Societies. *Science, New Series*, *312*(5781), 1767–1770.

Hicks, R., & Tingley, D. (2011). Causal Mediation Analysis. *The Stata Journal*, *11*(4), 1–14.

Imai, K., Keele, L., & Yamamoto, T. (2010). Identification, Inference and Sensitivity Analysis for Causal Mediation Effects. *Statistical Science*, *25*(1), 51–71.

Isaac, R.M., & Walker, J. M. (1988). Group Size Effects in Public Goods Provision: The Voluntary Contributions Mechanism. *The Quarterly Journal of Economics*, *103*(1), 179. doi:10.2307/1882648

Kamei, K., Putterman, L., & Tyran, J.-R. (2015). State or nature? Endogenous formal versus informal sanctions in the voluntary provision of public goods. *Experimental Economics*, *18*(1), 38–65. doi:10.1007/s10683-014-9405-0

Kesternich, M., Lange, A., & Sturm, B. (2014). The impact of burden sharing rules on the voluntary provision of public goods. *Journal of Economic Behavior & Organization*, *105*, 107–123. doi:10.1016/j.jebo.2014.04.024

Kesternich, M., Lange, A., & Sturm, B. (2018). On the performance of rule-based contribution schemes under endowment heterogeneity. *Experimental Economics*, *21*(1), 180–204. doi:10.1007/s10683-017-9535-2

Kosfeld, M., & Rustagi, D. (2015). Leader Punishment and Cooperation in Groups: Experimental Field Evidence from Commons Management in Ethiopia. *The American Economic Review*, *105*(2), 747–783.

Kroll, S., Cherry, T. L., & Shogren, J. F. (2007). Voting, Punishment, and Public Goods. *Economic Inquiry*, *45*(3), 557–570. doi:10.1111/j.1465-7295.2007.00028.x

Ledyard, J. O. (1995). Public Goods: A Survey of Experimental Research. *The Handbook of Experimental Economics*. https://authors.library.caltech.edu/80749/1/sswp861%20-%20revised.pdf. Accessed 28 May 2018

Magat, W., Krupnick, A. J., & Harrington, W. (1986). *Rules in the Making A Statistical Analysis of Regulatory Agency Behavior*. Baltimore: Taylor and Francis.

Masclet, D., Noussair, C., Tucker, S., & Villeval, M.-C. (2003). Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism. *The American Economic Review*, *93*(1), 366–380.

Messer, K. D., Zarghamee, H., Kaiser, H. M., & Schulze, W. D. (2007). New hope for the voluntary contributions mechanism: The effects of context. *Journal of Public Economics*, *91*(9), 1783–1799. doi:10.1016/j.jpubeco.2007.08.001

Murphy, J. J., & Stranlund, J. K. (2007). A laboratory investigation of compliance behavior under tradable emissions rights: Implications for targeted enforcement. *Journal of Environmental Economics and Management*, *53*(2), 196–212. doi:10.1016/j.jeem.2006.08.003

Nikiforakis, N., & Normann, H.-T. (2008). A comparative statics analysis of punishment in public-good experiments. *Experimental Economics*, *11*(4), 358–369. doi:10.1007/s10683-007-9171-3

Orzen, H. (2008). Fundraising through Competition: Evidence from the Lab, 38.

Ostrom, E. (1990). *Governing the commons: the evolution of institutions for collective action*. Cambridge ; New York: Cambridge University Press.

Ostrom, E., Walker, J., & Gardner, R. (1992). Covenants with and without a Sword: Self-Governance Is Possible. *American Political Science Review*, *86*(02), 404–417. doi:10.2307/1964229

Palfrey, T. R., & Rosenthal, H. (1991). Testing for effects of cheap talk in a public goods game with private information. *Games and Economic Behavior*, *3*(2), 183–220. doi:10.1016/0899-8256(91)90022-7

Palfrey, T., Rosenthal, H., & Roy, N. (2017). How cheap talk enhances efficiency in threshold public goods games. *Games and Economic Behavior*, *101*, 234–259. doi:10.1016/j.geb.2015.10.004

Putterman, L., Tyran, J.-R., & Kamei, K. (2011). Public goods and voting on formal sanction schemes. *Journal of Public Economics*, *95*(9–10), 1213–1222. doi:10.1016/j.jpubeco.2011.05.001

Qin, X., & Wang, S. (2013). Using an exogenous mechanism to examine efficient probabilistic punishment. *Journal of Economic Psychology*, *39*, 1–10. doi:10.1016/j.joep.2013.07.002

Rege, M., & Telle, K. (2004). The impact of social approval and framing on cooperation in public good situations. *Journal of Public Economics*, *88*(7–8), 1625–1644. doi:10.1016/S0047-2727(03)00021-5

Sefton, M., Shupp, R., & Walker, J. M. (2007). The Effect of Rewards and Sanctions in Provision of Public Goods. *Economic Inquiry*, *45*(4), 671–690. doi:10.1111/j.1465-7295.2007.00051.x

Shang, J., & Croson, R. (2009). A Field Experiment in Charitable Contribution: The Impact of Social Information on the Voluntary Provision of Public Goods. *The Economic Journal*, *119*(540), 1422–1439. doi:10.1111/j.1468-0297.2009.02267.x

Sutter, M., Haigner, S., & Kocher, M. G. (2010). Choosing the Carrot or the Stick? Endogenous Institutional Choice in Social Dilemma Situations. *The Review of Economic Studies*, *77*(4), 1540–1566.

Thibaut, J. W., & Walker, L. (1975). *Procedural justice: a psychological analysis*. Hillsdale, N.J. : New York: L. Erlbaum Associates ; distributed by the Halsted Press Division of Wiley.

Tyler, T. R. (1990). *Why people obey the law*. New Haven: Yale University Press.

Tyran, J.-R., & Feld, L. P. (2006). Achieving Compliance when Legal Sanctions are Non-deterrent. *Scandinavian Journal of Economics*, *108*(1), 135–156. doi:10.1111/j.1467-9442.2006.00444.x

Verba, S., Nie, N. H., & Kim, J. (1978). *Participation and political equality: a seven-nation comparison*. Cambridge, Eng. ; New York: Cambridge University Press.

Walker, J. M., Gardner, R., Herr, A., & Ostrom, E. (2000). Collective Choice in the Commons: Experimental Results on Proposed Allocation Rules and Votes. *The Economic Journal*, *110*(January), 212–234.

Yackee, S. W. (2006). Sweet-Talking the Fourth Branch: The Influence of Interest Group Comments on Federal Agency Rulemaking. *Journal of Public Administration Research and Theory*, *16*(1), 103–124. doi:10.1093/jopart/mui042

2. DO DIFFERENT EXTENSION APPROACHES AFFECT SMALLHOLDER FARMERS' WILLINGNESS-TO-PAY FOR NEW AGRICULTURAL TECHNOLOGIES? EXPERIMENTAL AUCTION RESULTS FROM TANZANIA

2.1 Introduction

Promoting the adoption and sustained use of new agricultural technologies is a major challenge for policymakers seeking to raise agricultural productivity and advance agricultural transformation (Barrett et al. 2017). In the absence of strong agricultural extension services, governments sometimes rely on donor-funded projects and non-governmental organizations (NGOs) to demonstrate and market new technologies in rural areas. While adoption and take-up of new technologies is often robust during the promotion and testing period of a new product, we frequently observe rapid dis-adoption and low take-up when the technologies are sold through traditional market channels (Moser and Barrett 2003; Hoffmann, Barrett and Just 2009; Bensch, Grimm and Peters 2015). The goal of this article is to better understand if the outreach efforts of NGO extension programs influence farmers' willingness to pay (WTP) for and awareness of new agricultural technologies. Focusing on seed for improved bean varieties and a new chemical seed treatment, Apron Star, this research seeks to answer the question: how do NGO lead-farmer extension programs influence the WTP for new technologies among Tanzanian bean producers?

Building on a previously implemented randomized controlled trial (RCT), we use a real auction approach to gauge technology demand and extend the existing literature in several important ways. First, we explicitly test whether lead-farmer technology demonstrations result in a higher producer WTP for seed for improved bean varieties and Apron Star compared to villages without such demonstrations. We also measure the effects of two extension approaches. One is a demonstration plot approach, where a lead-farmer maintains a plot in the village to educate other farmers about the use and benefits of a new technology. In the second approach, the lead-farmer maintains a demonstration plot and also distributes free trial packs of inputs to village farmers to test the new bean technologies on their own land. As lead-farmer methods become more popular (Kondylis, Mueller and Zhu 2017; Emerick and Dar 2017; Nakano et al. 2018; Whitfield et al. 2015), this research speaks to the benefits of incorporating more learning-by-doing (Foster and Rosenzweig 1995) into extension efforts (in this case, facilitated by the distribution of trial packs). Our second key contribution is to understand how farmers value agrodealer services in their communities, specifically the chemical treatment of seeds. We leverage differences in WTP between seeds pre-treated and self-treated with Apron Star to measure the demand for basic agricultural services both with and without lead-farmer demonstration plots and technology trial packs.

Working in partnership with the International Center for Tropical Agriculture (CIAT), fieldwork was conducted in August-September 2017 in the southern highlands region of Tanzania. We focus on the efforts of lead-farmers selected by an NGO, Farm Input Promotions-Africa (FIPS-Africa), who are using demonstration plots and the distribution of free small input trial packs to educate smallholders about improved bean and maize technologies. Drawing from a recently implemented RCT (Melkani and Mason 2018) focusing on improved bean technologies, we randomly selected 12 villages total in Mbeya Rural and Mbozi districts for inclusion in the study based on which lead-farmer extension approaches were employed in the village (6 demonstration plot only villages and 6 demonstration plot with trial packs villages). Additionally, we worked with FIPS-Africa to choose 6 control villages that met all the criteria for participation in the lead-farmer extension program but were not yet covered by FIPS-Africa due to funding constraints.

Within each village, 25 bean-growing households were randomly selected to participate in a household survey covering bean production history, engagement with lead-farmer extension activities, knowledge of improved bean varieties and Apron Star, and household demographics.

After completing the survey, households were invited to participate in a real Becker-DeGroot-Marschak (BDM) mechanism (Becker, DeGroot and Marschak 1964) for seed for two improved varieties of beans (Njano Uyole and Uyole 96) and the Apron Star seed treatment. Farmers received an endowment of 5,000 Tanzanian Shillings (Tsh), roughly US$2.25, to bid in the auction. Farmers placed bids on the following products for each bean variety: (i) 1 kg of certified, untreated bean seed; (ii) 1 kg of certified, untreated bean seed with a sachet of Apron Star (for farmers to use to self-treat the seed); and (iii) 1 kg of certified bean seed pre-treated with Apron Star. After all bids were placed, dice were rolled to select a single bid to be binding and to determine the random price.

Our findings on farmer WTP for seed for improved bean varieties and Apron Star seed treatment suggest that, on average, there is no effect of VBAA demonstration plots only or demonstration plots paired with trial packs when compared to the control group. A test for heterogeneous effects across the two districts suggests that in Mbozi district but not Mbeya Rural district, exposure to the demonstration plot only treatment decreases farmer WTP when compared to farmers in control villages or demonstration plots plus trial packs villages. We also find strong evidence that seed for improved bean varieties bundled with Apron Star is more valuable to smallholders in terms of WTP than seed not bundled with Apron Star. In addition, farmers are willing to pay a premium for seed pre-treated with Apron Star compared to untreated seed with a sachet of Apron Star for the farmer to apply him/herself.

The remainder of this paper is organized as follows. Section 2 reviews previous related research on agricultural extension and experimental auctions. Section 3 provides background on the agricultural technologies examined in this paper. In Section 4, we explain the conceptual framework for how exposure to different extension models may result in differences in farmer WTP for new technologies. Section 5 discusses the experimental design and sampling process. Section 6 presents and discusses results. Section 7 concludes by considering policy implications of this study and opportunities for further research.

2.2 Literature Review

2.2.1 Agricultural extension and delivery

Investment in agricultural extension systems is an important tool of policymakers to address the large gap in agricultural productivity observed in many developing regions. Beginning with the establishment of national agricultural advisory services in the 1950s, extension programs have been an important part of government budgets dedicated to agricultural growth and poverty reduction (Anderson 2008; Benin et al. 2011).

Anderson and Feder (2007) highlight several key types of agricultural extension models: Training and Visit (T&V), Fee-for-service, and Farmer Field Schools. T&V is the most common extension program where a formalized structure of in-house agricultural specialists and extension agents provide information and training to targeted villages on a set schedule (e.g., biweekly visits). This model is heavy on human capital, both for training and fieldwork, and has been implemented by national and local governments (Anderson and Crowder 2000). T&V model often involves working with lead-farmers who were successful early adopters of new agricultural technologies and management practices and would be able to share information with others (Aker 2011). Lead-farmers are then expected to train other farmers in their area about the use of new technologies or management practices. Fee-for-service models, often operated by private firms or public-private partnerships, focus on providing more specialized information or services to groups of farmers who pay. Farmer Field Schools (FFSs) were originally designed to promote integrated pest management in Southeast Asia and are now a widely used approach that brings together groups of farmers for multiple days to facilitate general information about agriculture, agronomy, and management practices (Aker 2011). FFS programs often engage lead-farmers as facilitators within their community to try to capture some benefits of social learning and heightened information transfer when farmers can learn from and more closely identify with one of their peers (Davis et al. 2012).[14]

Previous work on the effects of agricultural extension programs largely focuses on how program design affects outcomes related to technology adoption and uptake. Results on different extension modalities are mixed. Pan, Smith, and Sulaiman (2018) use a regression-discontinuity design to evaluate an NGO's lead-farmer extension program on agricultural technology adoption and food security among female smallholder farmers in Uganda. The program increases the probability that individuals adopt improved soil fertility management practices and improved seeds marketed through the NGO.[15] When looking at all improved varieties, however, the authors find little evidence that the program increases adoption. Smallholders in the treatment group experience a corresponding increase in food security measures, which the authors argue occurs through the adoption of improved production practices stimulated by the lead-farmer extension program.

[14] FFS do not focus primarily on the distribution of new agricultural technologies or even inputs associated with them. Instead the focus is broadly on general agricultural knowledge and practices.
[15] Improved soil fertility management practices whose adoption was increased include the application of animal manure, intercropping, crop rotation, and irrigation (Pan et al. 2018).

Leveraging an existing national lead-farmer program in Mozambique, Kondylis, Mueller, and Zhu (2017) evaluate the adoption consequences of adding a direct training module on sustainable land management (SLM) for some lead-farmers. Their study finds that additional direct training significantly increases SLM adoption among lead-farmers, but not among end-user farmers in treatment villages.[16] Unlike our study, there is no pure control group containing villages without lead-farmers so their findings do not speak to the overall impact of additional training. Emerick and Dar (2017) show that the addition of farmer field days to a lead-farmer extension model in India increases end-user farmer uptake of an improved variety of seed by 12 percentage points.

Nakano et al. (2018) evaluate the effects of a farmer-led training program on three distinct groups of rice producers in Tanzania: lead-farmers, farmers trained by lead-farmers, and end-user farmers in the village. Covering the adoption of improved seed, chemical fertilizer, and better management practices, a difference-in-differences analysis reveals that there were immediate positive effects of training on adoption among trained farmers. In addition, within three to four years, the new technologies also diffused to untrained (end-user) farmers in the village. These results are suggestive of the important time dimension of agricultural extension and training. It can take multiple production seasons for information about the new technologies to be collected and utilized by farmers not directly involved in the initial intervention.

There is a related group of studies that focuses on how the selection and incentive mechanisms behind lead-farmer extension programs can influence outcomes (Beaman et al. 2015; BenYishay and Mobarak 2015; Emerick and Dar 2017). For example, evidence from Malawi suggests that bypassing extension agents and choosing lead-farmers to demonstrate a new

[16] We use the term end-user farmer to refer to farmers in the village who are not trained as a lead-farmer and would be expected to learn or benefit from the application and use of a new agricultural technology.

technology via social network theory increases the adoption of an improved planting technique by 3 percentage points on average (Beaman et al. 2015). Emerick and Dar (2017) find no effect of lead-farmer selection methods on end-user adoption rates when comparing lead-farmers selected by local leaders to those selected by the community in India. Compensating lead-farmers can also increase effort exerted in communicating about new technologies across the village, further boosting adoption rates among end-user farmers in Malawi (BenYishay and Mobarak 2015).

2.2.2 WTP auctions in developing countries

Experimental auctions are widely used in the field as a tool to elicit individuals' valuations for goods and services. Real auctions, where bids are binding and money is exchanged for goods and services, can easily be conducted in the field to avoid the problem of hypothetical bias while capturing heterogeneity in valuations for a sample of interest (Lusk and Shogren 2007). Early applications of auction mechanisms in the developing country context focused on estimating consumer WTP for a myriad of products: e.g., certified baby food (Masters and Sanogo 2002), bed-nets (Hoffmann et al. 2009; Dupas 2014), and fortified maize meal (De Groote, Kimenju and Morawetz 2011).

Most relevant to our work is when these studies are conducted in the context of agricultural extension or outreach programs. De Groote et al. (2014) estimate the WTP of rural consumers in Tanzania for biofortified maize flour. While they find that consumers are willing to pay a significant price premium for the improved product relative to unfortified maize flour, they find no evidence that the extension program designed to promote the biofortified crop had any impact on individual WTP.

Recent research using experimental auctions in developing countries has evolved to also include producer WTP for improved inputs and agricultural services. Examples of auction studies include for laser land levelers (Lybbert et al. 2013), seed for improved crop varieties (Waldman, Kerr and Isaacs 2014), and safety equipment for chemical application (Goeb 2017). Similar to our study, Waldman, Kerr, and Isaacs (2014) estimate production preferences of Rwandan farmers for a common bean variety. In a similar vein to an extension program, some producers in their study are exposed to an on-farm trial where they grow all improved bean varieties in a demonstration plot prior to participating in experimental auctions. This is similar to the training that many lead-farmers are provided before being sent into the field. Farmers who participated in the on-farm trial were found to offer lower bids on average than farmers who only received yield information. This suggests that increased experience about a new technology may lower WTP as more information is obtained that is best observed in practice (e.g., days to maturity, weeding requirements).

2.3 Background

In this section we describe the improved bean technologies introduced to farmers in this study. We also describe the lead-farmer extension program being implemented in our treatment villages.

2.3.1 Improved varieties of beans in Tanzania

Uyole 96 is an improved bean variety released by CIAT and the Agricultural Research Institute Uyole (ARI-Uyole) in 1996. Traditional breeding methods were used to produce this line of large, dark red kidney beans from local cultivars that are used for both household consumption and sale as a cash crop (Hillocks et al. 2006). Consistently cited for high yields, Uyole 96 is tolerant to several common bean diseases including bean rust, asochyta, and Bean Common Mosaic Virus (BCMV) (Muhamba et al. 2013). Njano Uyole is a more recent bean variety released in 2008 by

ARI-Uyole. Njano Uyole is a medium size yellow bean that is tolerant to common bacterial blight (CBB), Alternaria leaf spot (ALS), halo blight, and root rot (Muhamba et al. 2013). Njano Uyole is also high yielding and farmers cite the relatively high market price and quality for cooking as positive characteristics (Em et al. 2013). In addition to the color and size of the two improved bean varieties, there are also other significant differences between them. Uyole 96 (Njano Uyole) takes 84 days (88 days) to mature, requires 36-40 kg (26-28 kg) of seed per acre, and has an expected yield of 480-1000 kg (600-1200 kg) of dry beans per acre.

These technologies allow us to investigate potential heterogeneous effects of the extension treatments between the two bean varieties. Furthermore, having two improved varieties with distinct production and consumption characteristics will increase the likelihood that bean growers would be willing to purchase at least one of the products in an experimental auction. Both varieties share elements of improved yield and disease tolerance but vary on elements of color and taste, which may be important for households that consume some portion of their farm output. An added benefit of these technologies is that they represent local or domestic innovation. Uyole 96 and Njano Uyole were both developed in consultation with domestic researchers at ARI-Uyole and require similar management practices to common local varieties of beans. Thus, the technologies may offer improved value without the adoption of complementary improved inputs or management practices, making both varieties a useful target for education by the lead-farmer extension program.

2.3.2 Apron Star seed treatment in Tanzania

Complementing the improved seed varieties, this article also looks at the introduction of a chemical seed treatment developed and commercialized by Syngenta and marketed in Tanzania under the name Apron Star. Billed as a fungicide-insecticide treatment, Apron Star is a chemical mixture that can be applied to bean (and maize) seed to control mildew, protect against early season insects (e.g., control sucking pests for 30 days after planting), and to prevent soil borne diseases (Syngenta 2017). Farmers can choose from several application methods including direct application as a dry dust, dry application to wet seed, or application to seeds as a water-based slurry. This means that the small-scale farmer at home can easily apply Apron Star to seed, using resources readily available on hand. Syngenta recommends using 5 g of Apron Star for every 2 kg of seed to achieve optimal results (Syngenta 2017).

While Apron Star is a novel technology in Tanzania, seed treatments and dressings have a long history of use in sub-Saharan Africa, promising increased yields through reduced risk of diseases and pests (Gibson 1953). An added benefit of these chemical treatments is that they do not incentivize the reduction of genetic diversity in cropped bean varieties as they can be applied to farmers' most preferred seed variety without requiring selection for particular genetic traits (Trutmann, Paul, and Cishabayo 1992).

In the southern highlands region of Tanzania, there is little evidence that seed treatments are a commonly used agricultural technology, especially in the production of legumes, despite the southern highlands being an important maize and bean-growing area in the country. In 2016, extension officers in the Mbeya region of Tanzania began recommending seed treatments similar to Apron Star for bean planting but there has not been widespread adoption of these technologies. Additionally, some agribusiness firms test and pilot new technologies in the Mbeya region before releasing them to the rest of the country. This raises the possibility that some sample farmers may have been exposed to similar seed treatment technologies prior to this study.

2.3.3 Village-Based Agricultural Advisors

FIPS-Africa is an NGO focused on improving food security and farmer incomes by making improved agricultural inputs and practices accessible to small-scale farmers. Founded in Kenya in 1990, FIPS has expanded their portfolio of improved agricultural inputs from inorganic fertilizer to include seed for improved crop varieties, pesticides, herbicides, and fungicides (Blackie and Albright 2005). Using a network of Village-Based Agricultural Advisors (VBAAs), FIPS provides extension services through technology demonstration plots and the distribution of free small trial packs of improved inputs to end-user farmers for them to test on their own plots.

VBAAs operate in their local community in Tanzania and are selected by their fellow community members based on farming experience, record keeping, communication skills, willingness to follow up with FIPS, and interest in becoming an agricultural input supplier.[17] Similar to other lead-farmer extension programs, we would expect this model to result in lead-farmers who have more experience with new technologies and are likely to be more educated than the average farmer in the village (Anderson and Feder 2007; Kondylis, Mueller, and Zhu 2017). VBAAs are all volunteers and are not paid employees of FIPS-Africa or other collaborating partners. Each VBAA serves as the primary point of contact between his/her village and the external public and private research institutes or firms who are interested in piloting or marketing new agricultural technologies. VBAAs' responsibilities may include providing information to other farmers in their community, establishing and maintaining one or more demonstration plots, and even distributing samples of agricultural technologies (e.g., small seed packs) to other farmers.

We focus on a sample of villages who have been randomly assigned to different VBAA outreach models as part of an RCT focusing on bidirectional learning and extension delivery

---

[17] FIPS also provides training in small business development and supports VBAAs to become registered fertilizer and seed dealers if they express interest.

(Melkani and Mason 2018). In the first treatment, VBAAs are assigned to conduct a demonstration plot for improved bean technologies in their village. The VBAA is provided necessary inputs and training to maintain the plot where they can plant traditional local varieties next to the new technologies for comparison.[18] In the second treatment, VBAAs also maintain the same demonstration plot but receive resources to deliver small bean input trial packs (100 g) to smallholder farmers in their village.[19] Prior to implementing the program in the village, all VBAAs in this study participated in an intensive direct training module on participatory extension methods and participatory learning (Melkani and Mason 2018).

2.4 Conceptual framework and hypotheses

We first seek to determine if there is an effect of VBAA extension and demonstration activities on farmer WTP for new agricultural technologies. There are two primary mechanisms through which the extension treatments might affect WTP: (i) learning-by-doing; and (ii) social learning or learning-from-others (Foster and Rosenzweig 1995; Munshi 2004; Bandiera and Rasul 2006; Conley and Udry 2010).[20] First, demonstration plots provide information to farmers who are unfamiliar with new technologies because the farmers are able to observe the production process and the new technologies in direct comparison with traditional production methods throughout the

---

[18] The bean demonstration plots included 16 sub-plots that were 12.5 m^2 in size (Melkani and Mason 2018). Each demonstration plot featured three improved varieties of bean seed (Uyole 96, Njano Uyole, Uyole 03) and a local variety selected by farmers. Four sub-plots were dedicated to each bean variety and were planted as follows: (i) untreated seed; (ii) seed treated with Apron Star; (iii) untreated seed with chemical fertilizer; (iv) seed treated with Apron Star and with chemical fertilizer applied (Melkani and Mason 2018).

[19] Trial packs were provided to FIPS VBAAs for distribution to 150 farmers in each village. Each trial pack contained 100g of seed for each of the following: an improved variety of bean seed both pre-treated with Apron Star and untreated, and a local variety of bean seed both pre-treated with Apron Star and untreated. This allowed for the comparison of improved vs. local bean varieties as well as Apron Star vs. untreated seed. 100g of bean seed is enough to plant a 2.5m x 5m plot.

[20] The literature on learning-by-doing and social learning broadly focuses on technology adoption as the key outcome of interest. This is different from our WTP measure because adoption occurs at a fixed market price. We would expect WTP to be highly correlated with technology adoption in the marketplace.

growing season. Additionally, a VBAA may involve other farmers in the village in the preparation or planting of the demonstration plot or other activities (e.g., weeding, harvesting) that provide hands-on experience with a new technology. Second, the distribution of free small input trial packs would allow some producers in the village to gain information about the performance of a new technology on their own land. Similar to the field days tested in Emerick and Dar (2017), provision of trial packs to village farmers significantly increases the probability that a farmer unconnected with the VBAA has a member of their social network who can provide information about the technology. This would also allow individuals to aggregate the experiences of different producers facing different production conditions (e.g., input application, soil conditions, etc.) and gain a more complete view of the technology.

Conditional on a new technology being profitable, we would expect that the acquisition of more information would increase farmers' WTP for improved bean seed technologies when compared to farmers with no access to new information, *ceteris paribus*. Comparing the two VBAA treatments however, we might expect WTP among farmers in the village with a demonstration plot paired with trial pack distribution to be higher than for a demonstration plot alone. Considering the same profitable technology, having more information about the performance across heterogeneous peer farmers (Munshi 2004; Magnan et al. 2015; Tjernström 2015; Emerick and Dar 2017) would increase WTP.

We also use the marketing of the Apron Star seed treatment technology to identify if farmers value self-treated and pre-treated seed differently. Once treated with Apron Star, seed can be stored for up to one year before being planted. Apron Star is a new technology to farmers in this region and there is very little local expertise about how to correctly apply and use the chemical seed treatment. An individual farmer would not necessarily know how to apply Apron Star unless they carefully read the Syngenta instructions, or they are taught how to apply it by the VBAA or a fellow farmer. Because of this unfamiliarity with a costly technology, we might expect farmers would be willing to pay a premium for pre-treated bean seed to ensure that the application was done properly. On the other hand, application of the Apron Star seed treatment is a relatively simple process that can be accomplished with readily available household tools. Self-treatment of the bean seed would be a way for farmers to reduce the input costs for the household and would be a reason to expect a small differential in WTP between the two different products.

Finally, we expect exposure to the two extension treatments to increase end-user farmer awareness of the technologies promoted via these programs. Similar to our expectations over the effects of treatment on WTP, we would expect the treatment that reaches more end-user farmers, demonstration plots paired with trial packs, to have a larger increase in end-user awareness of the new technologies than the demonstration plot alone. Knowledge is an important output measure because it reflects the ability of extension outreach to raise awareness about productivity enhancing bean technologies, even if the information does not necessarily change farmers valuation of those technologies.

2.5 Methods

In this research, we are interested in whether the type of lead-farmer extension program employed in a village affects small-scale farmers' WTP for improved bean technologies. We also want to understand how farmers might value the provision of local agricultural services – in this case pre-treatment of the bean seed with Apron Star. To address these issues, we conducted experimental auctions with small-scale bean farmers in Tanzania living in a set of villages targeted by the FIPS

VBAA lead-farmer extension programs as well as with a set of control villages who did not have a VBAA in the village.

2.5.1 Village selection

This study took place in 18 villages in August-September 2017 in the southern highlands of Tanzania. Villages were selected for the experimental auctions from an ongoing RCT training lead-farmers in bean agronomy and participatory extension approaches (Melkani and Mason 2018). Melkani and Mason (2018) focus on evaluating the effects of different extension approaches on VBAA-level outcomes and employ a pair-wise Mahalanobis matching algorithm (Bruhn and McKenzie 2009) based on observable, pre-intervention VBAA characteristics to increase balance among two treatment groups: demonstration plots only and demonstration plots combined with trial packs.

Using preliminary compliance data (Melkani and Mason 2018), we randomly selected 12 villages - 6 matched VBAA treatment pairs - to conduct the experimental auctions. Treatment pairs were equally divided across two administrative districts - Mbeya Rural and Mbozi. Additionally, we worked with FIPS-Africa to choose 6 control villages that met all the criteria for participation in the lead-farmer extension program but were not covered by FIPS-Africa due to funding constraints.[21] Control villages do not have a FIPS-Africa VBAA in the village. We are aware of no information distribution or training efforts on improved bean technologies having been conducted in these villages. Our final sample includes 18 auction villages equally divided into two treatment groups and a control group.

[21] However, because the original RCT design did not include a pure randomized control group, there remains the possibility of some selection bias in our identified control villages. This bias is generated because FIPS-Africa did not initially select these villages for inclusion in the VBAA program. While we argue that these villages were likely excluded due to overall funding constraints, there could also be unobservable factors that motivated the decision.

2.5.2 Farmer selection

Within each village, we randomly selected 25 bean growing households to participate in a bean production survey and experimental auction. Each village maintains a roster of current households and their members residing in the village and we used the following protocol to ensure random sampling. Upon arrival in the village, we met with the village chairperson, members of the village council, and, when available, the local agricultural extension agent and/or VBAA to ensure we had a complete and correct roster of current village households.[22] From the list of current households, we then worked with village leadership to identify all of the bean growing households in the village. For most villages this was a straightforward process as the village maintained agricultural records for each household by crop in addition to identifying information. We then rolled a six-sided die to determine the random start point for sampling of the bean growing households. In each village we selected 25 households for interview/auction participation with an additional 10 selected using the same process to serve as replacement households in the event that a selected household was not available or declined to participate in the study. We calculated the sampling interval by dividing the total number of bean growing households in the village by 25, the number that we wanted to survey.

2.5.3 Survey and experimental mechanism

In the experimental auctions, we used a BDM mechanism to collect participant bids for each bean seed product. Similar to a second-price auction, the BDM mechanism provides incentives such that the optimal strategy is for an individual to bid his/her true valuation of a given product (Lusk and Shogren 2007; De Groote et al. 2011). In the BDM mechanism, an individual wins the auction

[22] This process often involved removing households that were no longer in the village due to relocation or death in recent months.

and receives the product when his/her bid is higher than a randomly drawn price, but the farmer only pays the amount of the drawn price.

Working with implementing partners at CIAT and ARI-Uyole, we prepared six distinct products for use in the experimental auctions. For both improved bean varieties, Uyole 96 and Njano Uyole, we offered 1 kg of certified untreated seed, 1 kg of certified untreated seed with a 5 g sachet of Apron Star, and 1 kg of certified seed pre-treated with Apron Star. Seeds used in this experiment were purchased from certified growers affiliated with ARI-Uyole. Agronomists at ARI-Uyole treated the seeds with Apron Star following Syngenta recommendations using a wet slurry before packaging the seeds. All products were labeled and packaged in transparent bags to allow farmers the opportunity to inspect the product and observe the color and/or quality of the seeds presented to them.

The following steps were followed for all auction households participating in this study. First, subjects were read a statement of informed consent before completing a basic survey covering household characteristics, the past four seasons of bean production (the major and minor seasons of the previous two agricultural years – 2015/16 and 2016/17), and knowledge of improved bean technologies. The survey took approximately 45 minutes to complete for each household and enumerators recorded responses using tablets.

Second, enumerators read a statement of informed consent to each participant explaining confidentiality of bids and their rights as participants in the auction. Participants were informed that they could end the experiment at any point in time. Enumerators used a standard auction script to ensure that all instructions and elements of the BDM were presented the same way.

Third, individuals participated in a practice BDM auction using a bar of soap. After walking through an example of how the BDM mechanism works and why bidding your true value is the optimal strategy, participants received an endowment of 1,200 Tsh (US$1 is roughly 2,200 Tsh) with which to place bids on the soap. Individuals were instructed to place bids in 100 Tsh increments - this is commonly the smallest denomination of pricing, especially in rural areas. Once the participant placed their bid, the random price was determined by rolling a 12-sided die to generate a price from 0 to 1,100 Tsh. If the soap was purchased, the farmer paid the random price and received the soap, keeping any change left over.

Fourth, farmers were introduced to the improved bean technologies and were read and shown descriptions of the improved bean varieties and the Apron Star seed treatment. Descriptions of the Apron Star product were taken from the Syngenta product packaging and descriptions of the improved varieties of bean seed were provided by CIAT. Individuals were also informed that bean seed could be pre-treated with Apron Star or the chemical could be applied at home. Individuals were also provided a copy of the product descriptions to keep and reference throughout the bidding process.

Fifth, each farmer rolled a six-sided die to determine the random order in which they would submit bids for the six bean seed products. The auction order was determined at the individual level, see Table 2B.1 for more details on the possible orderings. After the order was selected, farmers were given 5,000 Tsh to place bids in increments of 100 Tsh on the six different products.

Sixth, participants placed bids on all six products following the selected order. Once all bids were submitted to the enumerator, the participant rolled a 6-sided die to determine which product bid would be binding.

Seventh, the random price was then determined for the binding product, using dice rolls to generate a random price from 0 to 4,900 Tsh. If the seed was purchased, the farmer paid the random price and received the seed, keeping any change left over. If the farmer purchased a product with pre-treated seed or with a sachet of Apron Star, enumerators also provided them with safety information about the product and safe handling instructions. All auction participants received contact information of who to approach (e.g., their VBAA, extension agents, and ARI-Uyole staff) if they had any additional questions about the improved bean technologies used in the experiment.

Table 2.1: Summary statistics by treatment group

VARIABLES	All (N=435)			Demonstration Plot Only (N=147)			Demonstration Plot + Trial Packs (N=144)			Control (N=144)		
	mean	min	max	mean	min	max	mean	min	max	mean	min	max
Demographics												
Respondent characteristics												
Gender (1 = Male)	0.543	0	1	0.578	0	1	0.507	0	1	0.542	0	1
Age (years)	43.650	18	90	42.210	18	73	46.380	22	90	42.400	19	90
Education (years)	6.400	0	21	6.993	0	18	5.972	0	21	6.222	0	17
Served on Village Council (1 = Yes)	0.267	0	1	0.259	0	1	0.271	0	1	0.271	0	1
Household characteristics												
Total land owned (acres)	4.050	0	30	4.099	0	20	3.804	0	23	4.244	0	30
Total titled land (acres)	0.464	0	28	0.374	0	20	0.312	0	11	0.708	0	28
Share of titled land (%)	6.150	0	100	5.101	0	100	5.800	0	100	7.569	0	100
PPI Poverty Likelihood (%)	21.150	1	62.1	21.860	1	62.1	18.610	1	62.1	22.980	1	62.1
Asset Index Score	0.033	-2.579	10.6	0.009	-2.579	9.269	-0.112	-2.546	7.776	0.202	-2.512	10.6
Extension/ Technology Experience												
Heard of Uyole 96 (1 = Yes)	0.526	0	1	0.537	0	1	0.562	0	1	0.479	0	1
Heard of Njano Uyole (1 = Yes)	0.393	0	1	0.361	0	1	0.458	0	1	0.361	0	1
Heard of Apron Star (1 = Yes)	0.078	0	1	0.129	0	1	0.076	0	1	0.028	0	1
Used Uyole 96 (1 = Yes)	0.285	0	1	0.306	0	1	0.271	0	1	0.278	0	1
Used Njano Uyole (1 = Yes)	0.117	0	1	0.061	0	1	0.153	0	1	0.139	0	1
Used Apron Star (1 = Yes)	0.007	0	1	0.000	0	0	0.014	0	1	0.007	0	1
Know of VBAA in Village (1 = Yes)	0.345	0	1	0.469	0	1	0.542	0	1	0.021	0	1

Table 2.1 (cont'd)

VARIABLES	All (N=435)			Demonstration Plot Only (N=147)			Demonstration Plot + Trial Packs (N=144)			Control (N=144)		
	mean	min	max	mean	min	max	mean	min	max	mean	min	max
Correctly Identify VBAA Status (1 = Yes)	0.662	0	1	0.469	0	1	0.542	0	1	0.979	0	1
HH member attended demo plot (1 = Yes)	0.044	0	1	0.068	0	1	0.063	0	1	0.000	0	0
HH received trial pack of seed (1 = Yes)	0.097	0	1	0.048	0	1	0.243	0	1	0.000	0	0
2017 Major Season Production	**Producing HH (N=327)**			**Producing HH (N=108)**			**Producing HH (N=108)**			**Producing HH (N=111)**		
# of 2017 Major Season (MS) Plots	1.116	1	3	1.120	1	3	1.093	1	3	1.135	1	3
2017 MS Total Plot Area (acres)	0.961	0.25	6.178	1.007	0.25	4	0.938	0.25	6.178	0.938	0.25	4
2017 MS Total Bean Area (acres)	0.932	0.25	6.178	0.965	0.25	3	0.914	0.25	6.178	0.916	0.25	4
2017 MS Bean Harvest (KG)	177.3	0	1,440	185.6	7	980	151.2	0	1,440	194.7	12	1,440
2017 MS Bean Sold (KG)	87.440	0	1,440	87.140	0	840	82.710	0	1,440	92.340	0	1,440
2017 MS Bean Price Received (Tsh/KG)	1149	300	2,200	1214	500	2,200	1191	500	2,200	1050	300	2,200
2017 MS Intercropped Beans (1 = Yes)	0.147	0	1	0.120	0	1	0.102	0	1	0.216	0	1
2017 MS Improved Variety (1 = Yes)	0.196	0	1	0.213	0	1	0.241	0	1	0.135	0	1
2017 MS Applied Apron Star (1 = Yes)	0.003	0	1	0.000	0	0	0.000	0	0	0.009	0	1
2017 MS Inorganic Fertilizer (1 = Yes)	0.419	0	1	0.472	0	1	0.361	0	1	0.423	0	1
2017 MS Herbicide (1 = Yes)	0.073	0	1	0.028	0	1	0.120	0	1	0.072	0	1
2017 MS Pesticide (1 = Yes)	0.495	0	1	0.389	0	1	0.454	0	1	0.640	0	1

Table 2.1 (cont'd)

	All (N=435)			Demonstration Plot Only (N=147)			Demonstration Plot + Trial Packs (N=144)			Control (N=144)		
VARIABLES	mean	min	max	mean	min	max	mean	min	max	mean	min	max
2017 MS Fungicide (1 = Yes)	0.180	0	1	0.231	0	1	0.185	0	1	0.126	0	1
2017 MS Manure (1 = Yes)	0.018	0	1	0.009	0	1	0.009	0	1	0.036	0	1

2.5.4 Summary statistics

We present summary statistics for key demographic, bean technology, and 2017 bean production variables in Table 2.1. The sample is slightly more heavily weighted towards men (54%) with an average age of 44 years old and six years of formal education. Subjects were primarily small-scale farmers who owned four acres of land on average, very little of which was titled (6%). To estimate a measure of the likelihood that a household would be classified as poor, we used a set of 10 questions to construct the Progress Out of Poverty Index and related likelihood score calibrated for Tanzania (Schreiner 2016). In the full sample, there is an average poverty likelihood of 21% across all of the households surveyed. Finally, we construct an asset index to approximate a wealth measure using the types and quantity of different assets owned by the household (e.g., livestock, mobile phones, computers, transportation, etc.).[23]

We test for balance across the experimental treatments, drawing on the demographic characteristics we collected for each household. This is important because the original random treatment assignment for the extension program was based on the matched characteristics of the VBAAs serving the village and not the characteristics of local bean end-user farmers. We exclude variables describing knowledge of improved bean technologies and 2017 bean production activities because these variables may have been influenced by the village's VBAA treatment assignment. The results of the balance test show that the demographic characteristics are jointly insignificant in predicting treatment assignment (F-test p-value > 0.1) and we fail to reject the null hypothesis of joint orthogonality (Table 2.2). A few statistically significant differences emerge between some of the treatments. For example, individuals in the demonstration plot only villages are slightly more educated than those in control villages at 7.0 years compared to 6.2 ($p<0.05$).

[23] We use principal components analysis (PCA) to construct this index measure.

Individuals in the demonstration plot with trial packs villages are 4.0 years older ($p<0.10$) and are 4.4 percentage points less likely to be in poverty ($p<0.01$) on average than those in the control group. We find no statistically significant differences between individuals assigned to the demonstration plot only and demonstration plot with trial pack treatments.

Table 2.2: Regression-based balance test based on observable respondent and household characteristics

	(1)	(2)	(3)
	Demo vs. Trial	Demo vs. Control	Trial vs. Control
Age	-0.003	0.002	0.004*
	(0.002)	(0.002)	(0.002)
Education	0.017	0.026**	0.009
	(0.011)	(0.011)	(0.011)
Gender (1 = Male)	0.044	0.012	-0.025
	(0.064)	(0.065)	(0.066)
Council membership	-0.023	-0.039	-0.032
	(0.072)	(0.072)	(0.072)
Land Owned (acres)	0.005	0.013	0.006
	(0.012)	(0.014)	(0.012)
Land Titled (acres)	0.023	-0.018	-0.018
	(0.030)	(0.019)	(0.019)
Share of land titled (%)	-0.002	0.000	0.001
	(0.002)	(0.002)	(0.002)
PPI Poverty Likelihood (%)	0.003	-0.002	-0.005**
	(0.002)	(0.002)	(0.002)
Asset Index	-0.001	-0.026	-0.020
	(0.019)	(0.021)	(0.019)
Constant	0.426***	0.254	0.363**
	(0.155)	(0.161)	(0.157)
Observations	291	291	288
R-squared	0.048	0.032	0.048
F-Test		**p-values**	
H_0: Joint Orthogonality	0.119	0.416	0.126

Note: OLS regressions. There are three treatment groups in this analysis: demonstration plot only (DP), demonstration plot with trial packs (DPTP), and the control. Dependent variables are binary variables indicating treatment status. In column (1) the dependent variable is equal to one if the individual is in the DP treatment and zero for DPTP. In column (2) the dependent variable is equal to one if the household is in the DP treatment and zero for control. In column (3) the dependent variable is equal to one if the household is in the DPTP and zero for control. Standard errors in parentheses. Significance is denoted by: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. We also present p-values at the bottom of each column for the joint test of orthogonality.

Summary results from the experimental auctions are presented in Table 2.3. We see that individual bids ranged from 200-5,000 Tsh across the 6 products up for auction. No participants submitted a bid of zero, suggesting that all individuals had some value or had some use for the bean production technologies being auctioned. This also increases our confidence in the sampling strategy for identifying bean growing households (assuming non-bean growing households are more likely to bid zero). The auction endowment was 5,000 Tsh, however we do observe at least one individual bidding the full value of the endowment for each product included in the analysis.[24] Turning to mean WTP for each product, we find that the control group consistently has the highest average WTP for each product. This suggests that the effect of either treatment on WTP is likely to be zero or negative, but we need to control for key sources of individual and geographic heterogeneity to better understand these differences.

[24] We also estimate Tobit models to account for this truncated bid data in the BDM. Results are substantively similar to the OLS results presented in this article and are available from the authors upon request.

Table 2.3: Experimental auction bid summary statistics by product bundle

	Demonstration Plot Only (N=147)			**Demonstration Plot + Trial Packs (N=144)**			**Control (N=144)**		
Product	mean	min	max	mean	min	max	mean	min	max
Untreated Uyole 96	2,498	500	5,000	2,317	200	5,000	2,529	400	5,000
Untreated Njano Uyole	2,492	500	5,000	2,438	200	5,000	2,567	300	5,000
Untreated Uyole 96 + Sachet	3,011	700	5,000	2,890	300	5,000	3,159	500	5,000
Untreated Njano Uyole + Sachet	3,048	500	5,000	2,980	300	5,000	3,200	500	5,000
Pre-treated Uyole 96	3,205	600	5,000	3,157	500	5,000	3,288	600	5,000
Pre-treated Njano Uyole	3,276	700	5,000	3,156	500	5,000	3,277	600	5,000

Note: All values are in Tsh per product. All seed purchased from the same certified supplier affiliated with ARI-Uyole's bean research department in August 2017. 2.5g of Apron Star is the Syngenta recommended application for 1kg of bean seed. All of the treated seed was treated using a slurry method where the treatment is applied to seeds wet and allowed to dry by ARI-Uyole staff.

2.6 Results

In the following section we present results for each of our main findings. First, we analyze the effects of the different VBAA extension treatments on farmer WTP for improved bean technologies. We look at main effects as well as heterogenous treatment effects by technology and district. We then investigate the effectiveness of the extension treatments in raising participant awareness of the technologies under consideration.

2.6.1 Effects of treatment on farmer WTP for improved technologies

Farmer WTP for an improved bean technology is our main outcome of interest. To estimate the effect of our extension treatments on farmer WTP, we use the following linear specification:

$$WTP_{ij} = \alpha + \beta_1 Demo_{ij} + \beta_2 Trial_{ij} + \boldsymbol{Prod}_j \boldsymbol{\gamma} + \boldsymbol{X}_i \boldsymbol{\delta} + \varepsilon_{ij} \quad (1)$$

where WTP_{ij} is the bid of farmer i for one of the six different improved bean seed technology products j in our real auction experiment. The parameter β_1 measures the treatment effect of having only demonstration plots in a village (*Demo* equals 1 in these villages and 0 otherwise), relative to the control group with no VBAA involvement. β_2 measures the treatment effect of having demonstration plots combined with the distribution of free input trial packs in a village (*Trial* equals 1 in these villages and 0 otherwise). $\boldsymbol{Prod}_j$ is a vector of indicator variables for the six different bean seed variety-Apron Star product bundles used in the auctions with untreated Uyole 96 serving as the reference category. $\boldsymbol{X}_i$ is a vector of household and respondent demographic characteristics as well as indicator variables for the auction bid order. Random treatment assignment was at the VBAA/village level, so we cluster all standard errors at the village level to allow for the correlation of treatment effects across households within a given village (Abadie et al. 2017). Because we have a relatively small number of clusters (18 villages) in our sample, it is necessary to present cluster-adjusted standard errors so as to avoid over-rejection of the null hypothesis (Colin Cameron and Miller 2015; Cameron, Gelbach and Miller 2008; Esarey and Menger 2018; Webb 2014). We implement a wild cluster bootstrap method (Wu 1986) to make this correction.

We find little evidence that the implementation of either VBAA extension treatment had an effect on farmer WTP for the bean seed technologies included in the experimental auctions. The treatment effects presented in column 1 of Table 2.4 for both the demonstration plot only and the demonstration plot with trial pack interventions are small in magnitude and not statistically different from zero.[25][26] While low power is a concern, the small magnitude of estimated effects suggest that even with increased power, there is unlikely to be a detectable effect. For example, the point estimate for the demonstration plot only treatment is 65.9 Tsh which would only represent a 2.2% change in WTP for untreated Uyole 96 and a smaller change for other products used in the auction.

We do, however, find evidence of large, statistically significant increases in WTP for products that include some version of Apron Star seed treatment. For example, take the case of Uyole 96. We estimate in column 1 that farmers are willing to pay 2937.8 Tsh for 1 kg of untreated seed. WTP increases by an estimated 571.8 Tsh for a 1 kg packet of Uyole 96 paired with an Apron Star sachet for a total valuation of 3,509.6 Tsh or a 19.5% increase in price. Similarly, we estimate an increase in WTP of 768.1 Tsh for 1 kg of Uyole 96 pre-treated with Apron Star, which is a 26.2% increase over the untreated product. The same pattern holds for the Njano Uyole products.

---

[25] See Table 2A.1 for results comparing only farmers in the original two treatment groups (excluding the control group).

[26] See Appendix 2B for a discussion of the effects of auction order on farmer WTP for improved bean products.

Table 2.4: OLS coefficient estimates for the effect of treatment on farmer WTP for improved bean products

	(1)	(2)	(3)
Demonstration Plot (DP)	-65.91	-104.06	377.27
	(0.798)	(0.638)	(0.310)
	[0.820]	[0.682]	[0.420]
Demonstration Plot + Trial Packs (DPTP)	-161.58	-147.08	0.31
	(0.560)	(0.463)	(0.999)
	[0.576]	[0.549]	[0.998]
Njano Uyole + Sachet	627.36***	627.36***	670.49***
	(0.000)	(0.000)	(0.000)
	[0.000]	[0.000]	[0.003]
Pretreated Njano Uyole	788.39***	788.39***	747.92
	(0.000)	(0.000)	(0.000)
	[0.000]	[0.000]	[0.130]
Untreated Njano Uyole	50.69	50.69	38.19
	(0.129)	(0.129)	(0.431)
	[0.119]	[0.119]	[0.422]
Uyole 96 + Sachet	571.84***	571.84***	629.86**
	(0.000)	(0.000)	(0.000)
	[0.000]	[0.000]	[0.017]
Pretreated Uyole 96	768.14***	768.14***	758.33**
	(0.000)	(0.000)	(0.000)
	[0.000]	[0.000]	[0.011]
Mbozi district (1 = Mbozi, 0 = Mbeya Rural)		377.18**	783.35*
		(0.025)	(0.030)
			[0.090]
DP x Mbozi			-839.53*
			(0.042)
			[0.095]
DPTP x Mbozi			-357.15
			(0.336)
			[0.471]
Education level (Respondent)		46.94**	40.63**
		(0.008)	(0.024)
		[0.030]	[0.031]
Bid order 2	-221.98	-184.01	-189.51
	(0.270)	(0.339)	(0.333)
	[0.262]	[0.344]	[0.336]
Bid order 3	-464.93**	-446.61**	-443.29**
	(0.018)	(0.023)	(0.023)
	[0.027]	[0.035]	[0.031]
Bid order 4	-476.53**	-449.28**	-486.67**
	(0.036)	(0.040)	(0.028)
	[0.050]	[0.045]	[0.0330]
Bid order 5	-577.35***	-536.51***	-503.73**
	(0.005)	(0.007)	(0.014)
	[0.004]	[0.009]	[0.013]
Bid order 6	-730.65***	-696.57***	-695.23***
	(0.005)	(0.008)	(0.008)
	[0.006]	[0.008]	[0.010]
Constant	2,937.80***	2,428.78***	2,256.53***
	(0.000)	(0.000)	(0.000)
	[0.000]	[0.009]	[0.000]

Table 2.4 (cont'd)

	(1)	(2)	(3)
Treatment x Item Interactions	No	No	Yes
Observations	2,610	2,610	2,610
R-squared	0.099	0.137	0.155

Notes: OLS regressions. Dependent variable is the farmer bid for a given product in Tanzanian Shillings (Tsh). Standard errors are clustered at the village level (18 clusters). Robust p-values in parentheses. In square brackets we report wild cluster p-values (Wu 1986) generated using boottest command in Stata 14 (Roodman 2015).
*** p<0.01, ** p<0.05, * p<0.1 for the wild cluster p-values respectively.

Column 2 shows that these inferences are robust to the inclusion of controls for the education level of the respondent and the district in which the respondent resides.[27] Across all model specifications, we find that farmers with a higher level of education have a higher WTP for improved bean technologies, as do farmers located in Mbozi district (relative to those located in Mbeya Rural district). An additional year of education is estimated to increase WTP for an improved bean product by 46.9 Tsh (p<0.01), or 1.9% of the estimated WTP for 1 kg of untreated Uyole 96. While small initially, for an individual who has completed through secondary school, this magnitude of the expected increase is comparable to adding an Apron Star sachet to one of the products. Living in Mbozi increases average WTP by 377.2 Tsh compared to living in Mbeya Rural, which is a 15.5% price premium for untreated Uyole 96. There are two possible reasons we might observe a significant district effect. First, it could be that the soil quality and production environment benefit more from these technologies (e.g. higher yield) and are valued by farmers. Second, households in Mbozi are a longer distance away from the ARI-Uyole research institute that produces and sells these improved lines of bean seed. The longer distance could reflect greater difficulty in acquiring the seed and thus a greater WTP for the product.

[27] With the exception of education level, all other demographic characteristics are jointly insignificant when included in the models (F-test p-value > 0.1). Based on the results of the joint F-test we drop gender, age, council membership, land holdings, land tenure, poverty likelihood, and the asset index score from models in Table 2.4. This is largely consistent with the findings of Waldman, Kerr, and Isaacs (2014) concerning the impacts of demographic characterics on WTP for different varieties of bean seed.

A test for heterogenous treatment effects between districts in column 3 reveals that the use of demonstration plots only in Mbozi district (but not Mbeya Rural district) has a negative and weakly significant effect on farmers' WTP with no corresponding effect when demonstration plots are paired with trial packs. After adjusting for the small number of clusters via bootstrapping however, this result is only marginally significant (p=0.095). This finding suggests that farmers may perceive the value of a technology differently when they are only exposed to information from the demonstration plot as opposed to being exposed to information both from the demonstration plot and the trials of multiple other farmers in their village.

Bean variety performance is highly susceptible to variation in climate and soil conditions, which can lead to different farmer preferences by their location (Waldman et al. 2014). Given some evidence of heterogenous treatment effects by location in the full sample, we next explore potential location-specific preferences using the same regression equation (equation 1) but restricting our sample to farmers residing in either Mbeya Rural or Mbozi district (Table 2.5).[28] Consistent with Table 2.4, we find no effect of the extension method used in Mbeya Rural district (column 1) on farmers' WTP for the new technology; and in Mbozi district, we again find a significant negative effect of the demonstration plot only treatment on WTP.

In columns 2 and 4 in Table 2.5, we check for heterogeneous treatment effects by product to determine if some product bundles are affected differently by the extension activities. In Mbeya Rural (column 2), we find that exposure to the demonstration plot with trial packs increases WTP by an average of 198.5 Tsh for 1 kg of Uyole 96 bean seed pre-treated with Apron Star. This suggests that having farmers experiment on their own plots with pre-treated Uyole 96 in addition to having the demonstration plot set up by the VBAA increased the perceived value of the product

[28] See Table 2A.2 for results comparing only farmers in the original two treatment groups (excluding the control group).

by farmers in Mbeya Rural. In contrast, in Mbozi (column 4), we find that exposure to the demonstration plot only treatment significantly decreases WTP by an average of 258.1 Tsh for 1 kg of Njano Uyole bean seed with 2.5 g of Apron Star for self-treatment. We find no evidence of heterogeneous treatment effects for the other eight treatment group-bean product pairs.

Table 2.5: Location specific OLS coefficient estimates for the effect of treatment on farmer WTP for improved bean products

VARIABLES	(1) Mbeya Rural	(2) Mbeya Rural	(3) Mbozi	(4) Mbozi
Demonstration Plot (DP)	295.00	232.96	-513.59**	-350.98
	(0.428)	(0.580)	(0.016)	(0.125)
	[0.548]	[0.653]	[0.032]	[0.178]
Demonstration Plot + Trial Packs (DPTP)	11.65	-112.10	-322.59	-263.52
	(0.971)	(0.740)	(0.140)	(0.266)
	[0.946]	[0.714]	[0.205]	[0.311]
Njano Uyole + Sachet	606.91***	554.17	647.71***	786.81***
	(0.000)	(0.001)	(0.000)	(0.000)
	[0.000]	[0.112]	[0.000]	[0.006]
Pre-treated Njano Uyole		644.44*	836.01***	851.39*
	(0.000)	(0.000)	(0.000)	(0.000)
	[0.000]	[0.069]	[0.000]	[0.068]
Untreated Njano Uyole	10.14	-55.56	91.06*	131.94
	(0.834)	(0.284)	(0.059)	(0.001)
	[0.838]	[0.371]	[0.069]	[0.101]
Uyole 96 + Sachet	499.54***	447.22*	643.81***	812.50*
	(0.000)	(0.000)	(0.000)	(0.000)
	[0.001]	[0.067]	[0.001]	[0.093]
Pre-treated Uyole 96	717.51***	611.11**	818.53***	905.56*
	(0.000)	(0.000)	(0.000)	(0.000)
	[0.000]	[0.021]	[0.000]	[0.056]
DP x NJ Sachet		18.06		-258.14***
		(0.934)		(0.000)
		[0.937]		[0.003]
DP x NJ Treated		194.44		-132.06
		(0.387)		(0.231)
		[0.429]		[0.309]
DP x NJ Untreated		44.44		-132.61
		(0.654)		(0.210)
		[0.722]		[0.355]
DP x U96 Sachet		-4.17		-231.83
		(0.980)		(0.156)
		[0.979]		[0.205]
DP x U96 Treated		119.44		-221.02
		(0.288)		(0.155)
		[0.339]		[0.210]
DPTP x NJ Sachet		138.98		-154.41
		(0.561)		(0.094)
		[0.598]		[0.139]

Table 2.5 (cont'd)

VARIABLES	(1) Mbeya Rural	(2) Mbeya Rural	(3) Mbozi	(4) Mbozi
DPTP x NJ Treated		93.91		92.27
		(0.452)		(0.490)
		[0.492]		[0.524]
DPTP x NJ Untreated		151.45		14.53
		(0.144)		(0.653)
		[0.194]		[0.715]
DPTP x U96 Sachet		159.63		-273.06
		(0.239)		(0.182)
		[0.277]		[0.247]
DPTP x U96 Treated		198.48**		-33.72
		(0.014)		(0.748)
		[0.030]		[0.758]
Education level	48.94**	48.94**	29.77	29.77
	(0.036)	(0.037)	(0.327)	(0.328)
	[0.027]	[0.027]	[0.326]	[0.326]
Constant	2,196.48***	2,258.69***	3,136.65***	3,061.47***
	(0.000)	(0.000)	(0.000)	(0.000)
	[0.002]	[0.005]	[0.000]	[0.002]
Dummy variable for bid order	Yes	Yes	Yes	Yes
Observations	1,302	1,302	1,308	1,308
R-squared	0.122	0.123	0.152	0.154

Notes: OLS regressions. Dependent variable is the farmer bid for a given product in Tanzanian Shillings (Tsh). Standard errors are clustered at the village level (9 clusters). Robust p-values in parentheses. In square brackets we report wild cluster p-values (Wu 1986) generated using boottest command in Stata 14 (Roodman 2015) and implementing weights via Webb (2014) to avoid spurious precision (Cameron et al. 2008).
*** p<0.01, ** p<0.05, * p<0.1 for the wild cluster p-values respectively.

2.6.2 Are farmers willing to pay for value-added agricultural services?

We next explore how the various combinations of improved bean seed technologies - variety and chemical seed treatment - influence farmer WTP. To do so, we use the following linear regression specification:

$$WTP_{ij} = \alpha + \beta_1 Demo_{ij} + \beta_2 Trial_{ij} + \gamma_1 Njano_j + \gamma_2 Sachet_j + \gamma_3 AST_j + \boldsymbol{X}_i \boldsymbol{\delta} + \varepsilon_{ij} \quad (2)$$

where WTP_{ij} is the bid of farmer i for improved bean product j in our real auction experiment. We again include binary treatment variables with the control group serving as the reference category. We are now interested in the attributes of each product. *Njano* is a binary variable equal to 1 if the bundle includes the Njano Uyole seed and 0 if it includes Uyole 96. Sachet is a binary variable equal to 1 if the bundle includes 5g of Apron Star for self-treatment of the bean seed and 0 otherwise. Finally, AST is a binary variable equal to 1 if the bundle contains pre-treated bean seed and 0 otherwise. We also again include $\boldsymbol{X_i}$, the vector of demographic characteristics and indicators for the auction bid order.[29] Standard errors are clustered at the village level and adjusted for the small number of clusters.

Table 2.6: OLS coefficient estimates for the effects of product attributes on farmer WTP

VARIABLES	(1)	(2)	(3)
Demonstration Plot (DP)	-65.91	-104.06	-81.21
	(0.798)	(0.637)	(0.712)
	[0.820]	[0.682]	[0.743]
Demonstration Plot + Trial Packs (DPTP)	-161.58	-147.08	-161.43
	(0.560)	(0.463)	(0.391)
	[0.576]	[0.549]	[0.447]
Seed variety (1=Njano Uyole, 0=Uyole 96)	42.15*	42.15*	22.80
	(0.091)	(0.092)	(0.465)
	[0.092]	[0.092]	[0.469]
Apron Star Sachet (1=Sachet included)	574.25***	574.25***	631.08***
	(0.000)	(0.000)	(0.000)
	[0.000]	[0.000]	[0.002]
Apron Star Pre-Treatment (1 = Pre-treated)	752.92***	752.92***	734.03***
	(0.000)	(0.000)	(0.000)
	[0.000]	[0.000]	[0.001]
Mbozi district		377.18**	377.18**
		(0.025)	(0.025)
		[0.030]	[0.030]
Education level (Respondent)		46.94***	46.94***
		(0.008)	(0.008)
		[0.009]	[0.009]
Constant	2,942.07***	2,433.05***	2,430.08***
	(0.000)	(0.000)	(0.000)
	[0.000]	[0.000]	[0.000]
Treatment x Item Attribute Interactions	No	No	Yes
Dummy variable for bid order	Yes	Yes	Yes
Observations	2,610	2,610	2,610
Hypothesis Test			
H_0: Apron Star Sachet = Apron Star Pre-Treatment	[0.000]	[0.000]	[0.051]

Notes: OLS regressions. Dependent variable is the farmer bid for a given product in Tanzanian Shillings (Tsh). Standard errors are clustered at the village level (18 clusters). Robust p-values in parentheses. In square brackets we report wild cluster p-values (Wu 1986) generated using boottest command in Stata 14 (Roodman 2015).
*** p<0.01, ** p<0.05, * p<0.1 for the wild cluster p-values respectively.

[29] Again, based on an F-test (p>0.1) we drop gender, age, council membership, land holdings, land tenure, poverty likelihood, and the asset index score from models in Table 2.5.

Small-scale farmers are willing to pay more for products including Apron Star and for the Njano Uyole seed variety (Table 2.6).[30] The results in column 1 suggest that farmers are willing to pay an average of 2942.1 Tsh for a 1 kg bag of untreated Uyole 96. There is a positive and significant premium associated with the Njano Uyole variety, however it is relatively small in magnitude at 42.2 Tsh or 1.4% of the mean WTP for the untreated Uyole 96 variety. Inclusion of the Apron Star sachet increases estimated WTP by an average of 574.3 Tsh or a 19.5% price increase relative to the base product (untreated Uyole 96). Finally, seed treated with Apron Star increases WTP by 752.9 Tsh, on average, which is a 25.6% increase over the base product.

We also test if there is a significant difference in producer WTP for seed pre-treated with Apron Star compared to seed bundled with a sachet of Apron Star that an individual could use to treat the seed on their own. Across all specifications in Table 2.6, we find that farmer WTP for pre-treated bean seed is higher than that of seed-Apron Star bundles requiring self-treatment. For example, in column 1 we estimate a 178.6 Tsh price premium for pre-treated seed over self-treated that is statistically significant at the 1% level. This premium is robust to several specifications, including the addition of controls (column 2) and the interaction of treatment with product characteristics (column 3).

Taken together, these results illustrate two key points. First, there is significant demand for improved bean technologies in the study districts, even among small-scale producers. Using the example of seed for improved bean varieties, the certified, untreated Uyole 96 and Njano Uyole can be purchased by any producer from ARI-Uyole for at 2,500Tsh/kg. The results above estimate mean WTP for certified, untreated Uyole 96 and Njano Uyole at 2,942.1 Tsh/kg and 2,984.2 Tsh/kg, respectively, which are approximately 18% above the wholesale price. Figures 2.1 and 2.2

[30] See Table 2A.3 for results comparing only farmers in the original two treatment groups (excluding the control group).

plot demand curves for the Uyole 96 and Njano Uyole products, respectively. In Figure 2.1, we illustrate that at the wholesale price of 2,500 Tsh/kg, 50% of farmers would be willing to purchase untreated Uyole 96 at this price, 68% would purchase untreated seed with an Apron Star sachet, and 75% would purchase seed pre-treated with Apron Star. Second, farmers have a high WTP for new technologies that do not necessarily require major investments in learning or effort. Consistent with our expectations we find a significant premium for pre-treated seeds. This could reflect a desire to forego some of the risk associated with a new technology (e.g. incorrectly applying Apron Star) or the costs associated with learning about the application of the new technology. Additionally, there is likely some opportunity cost of time in treating the seed at home. Among our sample, the average reported hourly wage rate for agricultural labor is 837.0 Tsh. Our estimated price premium for pre-treated seed over self-treated is 179.0 Tsh/kg or 21.4% of the hourly rate. Depending on how much seed is being treated, producers are probably rational to pay for the seed treatment.

Figure 2.1: Demand curves for Uyole 96

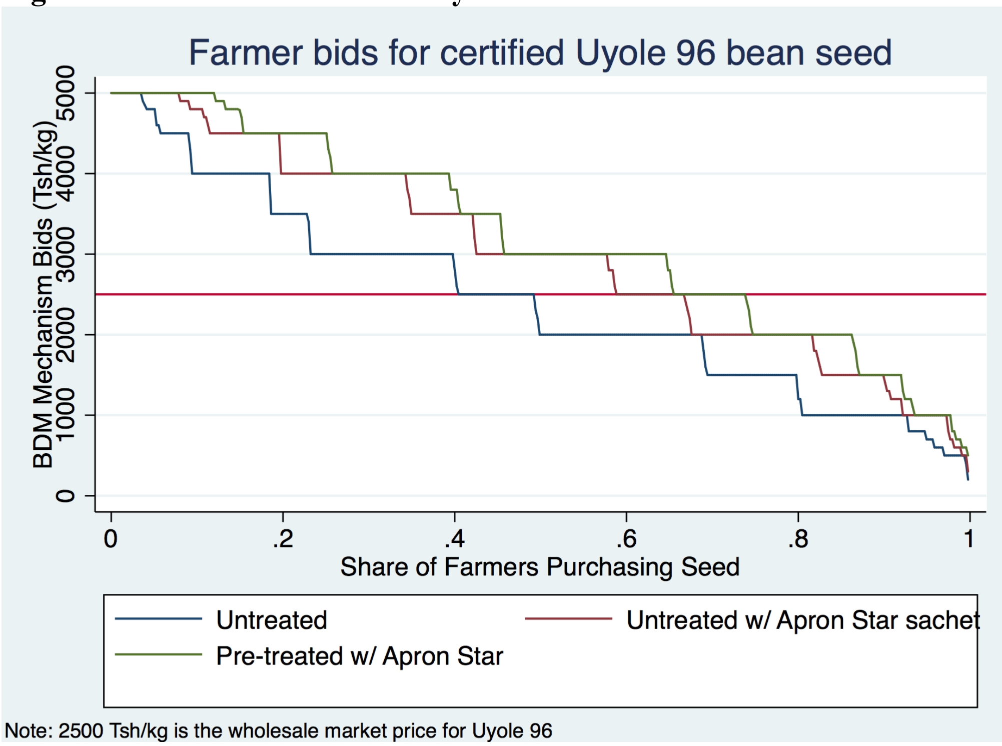
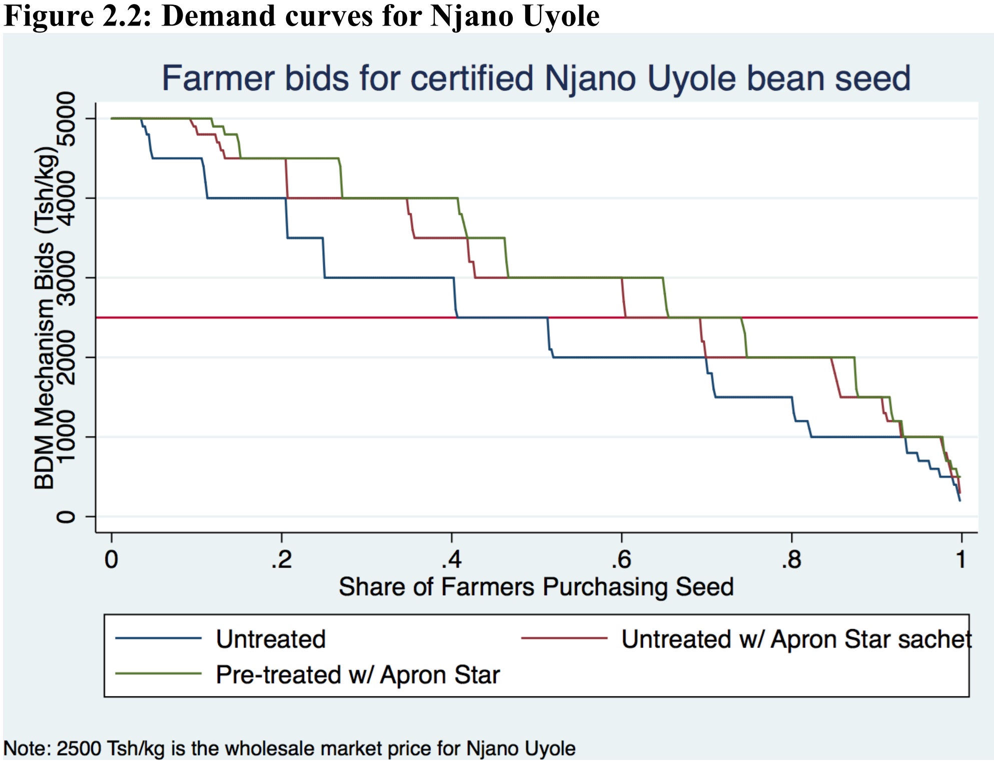


Figure 2.2: Demand curves for Njano Uyole



2.6.3 Are extension methods promoting learning and information transfer?

The delivery of extension services is expected to impact farmer awareness of and knowledge about new agricultural technologies. Conditional on expectations of profitability, this knowledge may or may not translate into differences in farmer WTP for a product. We test the learning mechanism using the following probit specification:

$$P(Knowledge_{ij} = 1 | Demo_{ij}, Trial_{ij}, \boldsymbol{X}_i) = \Phi(\alpha + \beta_1 Demo_{ij} + \beta_2 Trial_{ij} + \boldsymbol{X}_i \boldsymbol{\delta}) \quad (3)$$

where $Knowledge_{ij}$ is a binary measure equal to one if farmer i is aware of bean technology j and zero otherwise. We control for treatment status of the village as well as $\boldsymbol{X}_i$, a vector of demographic characteristics as defined above. We estimate separate equations for our two varieties of interest as well as for the Apron Star seed treatment.

Table 2.7: Effects of treatment on farmer knowledge of a bean technology

VARIABLES	(1) Uyole 96	(2) Njano Uyole	(3) Apron Star
Demonstration Plot (DP)	0.16	0.05	0.90***
	(0.724)	(0.818)	(0.000)
	[0.747]	[0.838]	[0.000]
Demonstration Plot + Trial Packs (DPTP)	0.31	0.34	0.54***
	(0.367)	(0.222)	(0.000)
	[0.430]	[0.252]	[0.008]
Age	-0.00	-0.01*	-0.00
	(0.258)	(0.069)	(0.722)
	[0.303]	[0.086]	[0.752]
Education	0.05**	-0.02	-0.04
	(0.025)	(0.263)	(0.349)
	[0.032]	[0.326]	[0.324]
Gender (1 = Male)	-0.25	-0.06	0.07
	(0.108)	(0.666)	(0.769)
	[0.109]	[0.664]	[0.770]
Served on Village Council (1 = Yes)	0.15	0.32	0.50*
	(0.385)	(0.123)	(0.051)
	[0.381]	[0.129]	[0.066]
Total land owned (acres)	-0.01	0.01	-0.01
	(0.421)	(0.524)	(0.822)
	[0.454]	[0.502]	[0.854]
Share of titled land (%)	0.01*	0.01*	-0.01**
	(0.023)	(0.055)	(0.042)
	[0.058]	[0.080]	[0.020]
PPI Poverty Likelihood (%)	0.00	0.00	-0.00
	(0.669)	(0.535)	(0.705)
	[0.676]	[0.580]	[0.698]
Asset Index Score	0.11***	0.03	0.05
	(0.005)	(0.324)	(0.343)
	[0.004]	[0.332]	[0.404]
Mbozi district	0.24	0.24	-0.35***
	(0.465)	(0.248)	(0.004)
	[0.455]	[0.253]	[0.005]
Constant	-0.26	-0.11	-1.52**
	(0.564)	(0.804)	(0.023)
	[0.565]	[0.811]	[0.052]
Observations	435	435	435
Hypothesis Testing			
H_0: DP = DPTP	[0.725]	[0.247]	[0.009]

Notes: Probit regressions. Dependent variable is equal to 1 if the farmer has ever heard of a specific bean technology and equal to zero otherwise. We look at two varietal technologies (Uyole 96 and Njano Uyole) and the Apron Star seed treatment. Standard errors are clustered at the village level (18 clusters). Robust p-values in parentheses. In square brackets we report wild cluster p-values (Wu 1986) generated using boottest command in Stata 14 (Roodman 2015). *** p<0.01, ** p<0.05, * p<0.1 respectively for the wild cluster p-values

We find mixed evidence on the effectiveness of the two extension treatments to increase end-user farmer knowledge about bean seed technologies. Columns 1 and 2 in Table 2.7 provide little support that exposure to a demonstration plot or a demonstration plot with trial packs increases farmer awareness of the improved varieties of beans used in the experimental auctions.[31] The point estimates on the extension treatment terms are insignificant, and indistinguishable from zero. The estimates are positive however, and we cannot rule out low power as the reason we can't detect a relatively small effect size. Another possible explanation for this finding is that these particular bean seed technologies have reached a market saturation point among farmers who would be interested in purchasing and planting an improved bean seed rather than a local variety. Uyole 96 was introduced by ARI-Uyole and CIAT in 1996, while Njano Uyole was introduced in 2008, but market availability likely took some time to materialize. This explanation seems less likely given that across the sample only 52.6% and 39.3% percent of farmers have heard of Uyole 96 and Njano Uyole varieties respectively. Demographic characteristics do appear to play a role in which farmers have heard of each improved variety. On average, farmers who are more educated, hold more assets, and have a higher share of titled land are more likely to have heard of Uyole 96. Younger farmers and those with more assets are more likely to have heard of Njano Uyole.

The Apron Star seed treatment is the newest technology employed by VBAAs, and potentially has the most room for information dissemination. In Table 2.7 column 3, we find strong evidence that both extension treatments increase farmer awareness of Apron Star compared to farmers in the control group. Exposure to the demonstration plot only treatment increases the probability that a farmer is aware of Apron Star by 11.0 percentage points on average. Farmers in

[31] See Table 2A.4 for results comparing only farmers in the original two treatment groups (excluding the control group).

the demonstration plot with trial pack treatment are 4.9 percentage points more likely to be aware of the new technology.[32]

These findings suggest that these extension programs do appear to be a viable mechanism to increase the awareness of producers about some technologies, especially those that may be very new to the market. The fact that our extension treatments increase awareness however, does not imply that there is a corresponding effect on WTP as our previous results have shown. One explanation for this is that information received from demonstration plots and neighbors can be a noisy signal with a significant amount of uncertainty surrounding the expected gains from uptake (Foster and Rosenzweig 1995; Suri 2011). Even in villages where trial packs were distributed, the small pack size was only enough to seed approximately 12.5 square meters (0.003 acres) with a given variety of seed. Given heterogeneity in plot types and soil quality, such small plots are unlikely to provide a clear signal about potential profitability even to farmers who experiment with the technology themselves. This could explain why awareness imperfectly translates into higher WTP (assuming the technology is profitable). A second explanation for finding a positive effect on awareness but not on WTP is the low level of general awareness of the technology. Prior to conducting the auction (but after the treatments were implemented), 12.9% of the demonstration plot only sample and 7.6% of the demo plot with trial pack sample were aware of Apron Star. While relatively higher than the 2.8% awareness rate in the control group, there may not have been enough individuals exposed to the training or information to drive an average difference in WTP.

[32] Note that these are average partial effects (APE) from the probit model for ease of interpretation. We use APEs because the coefficients from nonlinear-in-parameters models like probit indicate the *direction* of the effects of the independent variables (positive or negative) but must be scaled in order for their *magnitudes* to be interpretable. Probit APEs can be interpreted as the estimated average change in the dependent variable given a one unit increase in the independent variable of interest in the continuous independent variable case, and given a discrete change in the independent variable of interest in the binary independent variable case, holding other factors constant.

Additionally, we find that the effect of the demonstration plot only treatment on producer awareness of Apron Star is 6.1 percentage points greater (p-value = 0.009) than in the treatment paired with trial packs. This result is surprising given that both treatments included the same demonstration plot design. One possible explanation could be that community members were more engaged with the demonstration plot when it was the only source of information about the new technology. Farmers could have made repeated visits to observe differences in seasonal growth, where once they received a trail pack and conducted their own experiment there was little motivation to revisit the demonstration plot.

2.7 Conclusions and policy implications

Public and private sector extension programs often incorporate lead-farmer demonstration activities and trial packs in an effort to increase the number of opportunities farmers have to learn about technologies through both their own experience and the experiences of others. These mechanisms are expected to increase adoption and diffusion of the improved agricultural technologies they showcase, assuming there are profitability gains to be made by the farmers they target. Building on an RCT focused on varied farmer-led NGO extension approaches (Melkani and Mason 2018), this article leverages real auctions to test the effect of these extension approaches on farmers' WTP for improved bean technologies. The first treatment introduces new bean technologies to a village via a demonstration plot, where farmers can come to learn and observe the performance of new technologies. The second treatment combines the same demonstration plot with the distribution of trial packs of inputs in the village so that farmers can experiment with the technologies on their own land. We focus on seed for improved varieties (Uyole 96 and Njano Uyole) as well as a chemical seed treatment called Apron Star. Given the need to treat bean seeds with Apron Star, we also investigate how producers value the provision of basic agricultural services (application of seed treatment) by agricultural input suppliers; agricultural service provision in general is a potentially significant area of future demand and employment growth in rural areas. This research makes two main contributions to the literature. First, we evaluate the impacts on village farmers of adding trial packs, a significant source of learning-by-doing, to a farmer-led extension model. This complements recent work evaluating the impacts of different extension modalities on farmer outcomes (Kondylis et al. 2017; Emerick and Dar 2017). Second, we extend the application of WTP modules to an extension and supply-side framework, providing insight into the market viability of improved bean technologies in rural areas. Specifically, we focus on producer WTP for improved inputs into the bean production process rather than on consumption goods.

We find little evidence that exposure to either extension treatment significantly impacts farmer WTP for improved bean technologies, despite increasing awareness of the Apron Star seed treatment. In fact, when we look at location-specific effects, we find weak evidence that farmers exposed to the demonstration plot only treatment have a significantly *lower* WTP than farmers in the control group or demonstration plots plus trial packs treatment group. One explanation for this finding is that the new technologies may not represent a profitable improvement for producers, especially under specific agricultural and climate conditions. Because information from the demonstration plot does not translate perfectly to a farmers' own land, these imprecise signals could lead to a downward adjustment in WTP. The addition of trial packs and own experimentation on a farmer's own land seems to attenuate this marginally significant result. Another possible explanation has to do with access to new technologies in an auction context. We could be observing farmers who already have access to some improved technologies and are not willing to pay as much to procure new ones (e.g. a specific variety of bean seed). Similarly, producers in the control group – with no access to new technologies – might bid higher for improved technologies just for the opportunity to test them on their farms. Combined, these effects would make it more difficult to detect any treatment effects of extension services. We find no significant difference in WTP between the two extension treatment groups (demonstration plots only vs. demonstration plots plus trial packs). It is also important to note that these WTP results are in the context of low power due to the small number of clusters where experimental auctions were conducted. While we are able to adjust our standard errors for a small number of clusters using wild cluster bootstrap p-values for all regressions, a challenge of power still remains. This concern is alleviated to some extent by the fact that many of our point estimates of WTP are relatively small (<10% of the total effect size) and would be unlikely to be significant in the context of a much larger study.

There is evidence, however, that farmers are willing to pay for improved bean technologies. We find that producers in our sample are willing to pay a premium of 1.5% for Njano Uyole over Uyole 96, 19.5% for improved varieties with an Apron Star sachet, and 25.6% for improved varieties treated with Apron Star. Individual valuation for pre-treated seed is also 5.1% greater than for the same amount of seed plus Apron Star that would have to be applied by the farmer. Taken together, these results suggest there is significant demand for new technologies and even the provision of services among small-scale farmers.

Our experimental results point to the following policy implications. First, if lead-farmer extension efforts paired with demonstration plots or with trial packs do not raise farmer valuations of ostensibly beneficial technologies, they may be diverting critical resources away from the provision of traditional government extension services. This would be true even when activities are funded by donors or NGOs, if the government views private extension services as a substitute for traditional outreach efforts. Despite the close proximity of lead farmers both socially and geographically to other farmers in the village they are serving and the relatively low cost of recruiting participants, the quality of lead-farmer extension activities is extremely difficult to monitor which could influence their effectiveness. Pursuing public-private partnerships focused on capacity building by partnering with lead-farmers (e.g., Syngenta with Apron Star) and public extension may represent a viable path forward. Second, illustrates the continued demand and WTP for some improved agricultural technologies when offered in rural communities. Lead farmer extension services may instead represent a promising field laboratory to refine which technologies are commercially marketed to which regions. Third, we do find evidence that smallholder farmers are willing to pay a premium for the provision of relatively simple agricultural services. Providing value-added support services like seed treatment, not demonstration and education, might be the most important role of VBAAs moving forward.

This article also highlights several areas ripe for future research. More work needs to be undertaken to understand the effects of extension programs on farmer WTP for new technologies for multiple types of goods to understand how producers might respond in the marketplace once promotional programs end. With the continued proliferation of actors in the extension space (Anderson and Feder 2007), introducing auctions might allow us to compare the effects of extension outreach across private and public extension models. Furthermore, more work needs to be done to understand which agricultural services smallholder farmers would be willing to purchase from local agrodealers instead of performing themselves. Not only do these services potentially reduce required costs and expertise on the part of the household but they also represent a significant area for potential employment growth in the agricultural value chain serving smallholders.

APPENDICES

APPENDIX 2A: Robustness checks

Table 2A.1: Effect of treatment on WTP excluding the control group

VARIABLES	(1)	(2)	(3)	(4)
Demonstration Plot + Trial Packs (DPTP)	-109.51	-38.77	-52.82	-351.81*
	(0.481)	(0.762)	(0.655)	(0.087)
	[0.504]	[0.790]	[0.688]	[0.155]
Njano Uyole + Sachet	606.01***	606.01***	606.01***	550.00**
	(0.000)	(0.000)	(0.000)	(0.000)
	[0.000]	[0.000]	[0.000]	[0.011]
Pretreated Njano Uyole	808.42***	808.42***	808.42***	777.89**
	(0.000)	(0.000)	(0.000)	(0.000)
	[0.000]	[0.000]	[0.000]	[0.028]
Untreated Njano Uyole	56.87	56.87	56.87	-5.78
	(0.202)	(0.204)	(0.203)	(0.927)
	[0.194]	[0.194]	[0.194]	[0.845]
Uyole 96 + Sachet	543.13***	543.13***	543.13***	513.27**
	(0.000)	(0.000)	(0.000)	(0.000)
	[0.000]	[0.000]	[0.000]	[0.016]
Pretreated Uyole 96	772.99***	772.99***	772.99***	707.07**
	(0.000)	(0.000)	(0.000)	(0.000)
	[0.000]	[0.000]	[0.000]	[0.023]
DPTP x Njano Uyole + Sachet				113.19
				(0.434)
				[0.471]
DPTP x Pretreated Njano Uyole				61.69
				(0.649)
				[0.631]
DPTP x Untreated Njano Uyole				126.62
				(0.118)
				[0.128]
DPTP x Uyole 96 + Sachet				60.35
				(0.611)
				[0.633]
DPTP x Pretreated Uyole 96				133.20
				(0.168)
				[0.185]
Mbozi district		205.99	185.32	-26.17
		(0.196)	(0.178)	(0.884)
		[0.286]	[0.236]	[0.896]
DPTP x Mbozi district				427.55
				(0.082)
				[0.126]
Education (Respondent)		32.76	51.01**	46.16*
		(0.234)	(0.019)	(0.042)
		[0.242]	[0.030]	[0.059]
Age		-8.54		
		(0.176)		
		[0.226]		
Gender (1 = Male)		64.09		
		(0.561)		
		[0.597]		

Table 2A.1 (cont'd)

VARIABLES	(1)	(2)	(3)	(4)
Served on Village Council (1 = Yes)		153.16		
		(0.417)		
		[0.483]		
Total land owned (acres)		-1.17		
		(0.971)		
		[0.974]		
Share of titled land (%)		-0.17		
		(0.967)		
		[0.969]		
PPI Poverty Likelihood (%)		0.77		
		(0.881)		
		[0.879]		
Asset Index Score		-36.21		
		(0.500)		
		[0.543]		
Constant	2,909.00***	2,802.99***	2,435.10***	2,621.28***
	(0.000)	(0.000)	(0.000)	(0.000)
	[0.000]	[0.000]	[0.000]	[0.000]
Bid order 2	-77.18	-10.03	-60.88	-81.60
	(0.791)	(0.974)	(0.826)	(0.770)
	[0.786]	[0.980]	[0.813]	[0.752]
Bid order 3	-587.73**	-556.17**	-589.41**	-594.08**
	(0.026)	(0.044)	(0.017)	(0.015)
	[0.035]	[0.059]	[0.019]	[0.014]
Bid order 4	-430.20	-365.38	-386.95	-408.45
	(0.125)	(0.172)	(0.159)	(0.129)
	[0.138]	[0.190]	[0.166]	[0.141]
Bid order 5	-632.52**	-590.40**	-607.54**	-606.73**
	(0.028)	(0.043)	(0.026)	(0.031)
	[0.028]	[0.052]	[0.024]	[0.026]
Bid order 6	-914.21***	-848.48**	-862.84***	-843.41***
	(0.007)	(0.012)	(0.007)	(0.009)
	[0.009]	[0.013]	[0.009]	[0.011]
Observations	1,746	1,746	1,746	1,746
R-squared	0.128	0.161	0.151	0.158

Notes: OLS regressions. Dependent variable is the farmer bid for a given product in Tanzanian Shillings (Tsh). Standard errors are clustered at the village level (12 clusters). Robust p-values in parentheses. In square brackets we report wild cluster p-values (Wu 1986) generated using boottest command in Stata 14 (Roodman 2015).
*** p<0.01, ** p<0.05, * p<0.1 for the wild cluster p-values respectively.

Table 2A.2: Location specific treatment effects excluding the control group

VARIABLES	(1) Mbeya Rural	(2) Mbeya Rural	(3) Mbozi	(4) Mbozi
Demonstration Plot + Trial Packs (DPTP)	-293.10*	-354.81	203.51	99.96
	(0.075)	(0.214)	(0.246)	(0.552)
	[0.096]	[0.258]	[0.308]	[0.634]
Njano Uyole + Sachet	633.10***	572.22**	579.11***	528.67***
	(0.007)	(0.028)	(0.000)	(0.000)
	[0.008]	[0.082]	[0.001]	[0.265]
Pretreated Njano Uyole	788.28***	838.89**	828.42***	719.33***
	(0.001)	(0.011)	(0.000)	(0.000)
	[0.003]	[0.0735]	[0.002]	[0.051]
Untreated Njano Uyole	42.76	-11.11	70.89	-0.67
	(0.532)	(0.901)	(0.294)	(0.995)
	[0.505]	[0.823]	[0.282]	[0.983]
Uyole 96 + Sachet	525.52***	443.06**	560.62***	580.67***
	(0.004)	(0.035)	(0.000)	(0.000)
	[0.009]	[0.084]	[0.002]	[0.047]
Pretreated Uyole 96	770.34***	730.56***	775.62***	684.53***
	(0.000)	(0.001)	(0.000)	(0.002)
	[0.003]	[0.082]	[0.002]	[0.094]
DPTP x Njano Uyole + Sachet		120.93		103.73
		(0.683)		(0.280)
		[0.716]		[0.335]
DPTP x Pretreated Njano Uyole		-100.53		224.33*
		(0.688)		(0.095)
		[[0.697]		[0.104]
DPTP x Untreated Njano Uyole		107.00		147.15
		(0.408)		(0.197)
		[0.453]		[0.309]
DPTP x Uyole 96 + Sachet		163.79		-41.23
		(0.436)		(0.769)
		[0.509]		[0.731]
DPTP x Pretreated Uyole 96		79.03		187.30
		(0.549)		(0.239)
		[0.629]		[0.296]
Education (Respondent)	36.88	36.88	46.96	46.96
	(0.237)	(0.238)	(0.200)	(0.201)
	[0.147]	[0.147]	[0.221]	[0.221]
Bid order 2	131.15	131.15	-283.41	-283.41
	(0.680)	(0.681)	(0.561)	(0.562)
	[0.676]	[0.676]	[0.551]	[0.551]
Bid order 3	-346.18**	-346.18**	-834.96*	-834.96*
	(0.049)	(0.049)	(0.054)	(0.055)
	[0.940]	[0.940]	[0.079]	[0.079]
Bid order 4	-344.81	-344.81	-475.98	-475.98
	(0.383)	(0.384)	(0.170)	(0.171)
	[0.363]	[0.363]	[0.240]	[0.240]
Bid order 5	-635.14*	-635.14*	-580.99	-580.99
	(0.064)	(0.065)	(0.197)	(0.199)
	[0.051]	[0.051]	[0.206]	[0.206]
Bid order 6	-896.57**	-896.57**	-793.19	-793.19
	(0.020)	(0.020)	(0.138)	(0.139)
	[0.005]	[0.005]	[0.185]	[0.185]

Table 2A.2 (cont'd)

VARIABLES	(1) Mbeya Rural	(2) Mbeya Rural	(3) Mbozi	(4) Mbozi
Constant	2,573.34***	2,604.41***	2,600.89***	2,651.24***
	(0.000)	(0.000)	(0.002)	(0.002)
	[0.003]	[0.006]	[0.009]	[0.023]
Observations	870	870	876	876
R-squared	0.186	0.187	0.136	0.138

Notes: OLS regressions. Dependent variable is the farmer bid for a given product in Tanzanian Shillings (Tsh). Standard errors are clustered at the village level (6 clusters). Robust p-values in parentheses. In square brackets we report wild cluster p-values (Wu 1986) generated using boottest command in Stata 14 (Roodman 2015) and implementing weights via Webb (2014) to avoid spurious precision (Cameron et al. 2008).
*** p<0.01, ** p<0.05, * p<0.1 for the wild cluster p-values respectively.

Table 2A.3: WTP for product characteristics excluding the control group

	(1)	(2)	(3)	(4)	(5)
Demonstration Plot + Trial Packs (DPTP)	-109.51	-38.77	-52.82	-551.74*	-108.00
	(0.481)	(0.762)	(0.655)	(0.067)	(0.482)
	[0.504]	[0.790]	[0.688]	[0.115]	[0.561]
Seed variety (1=Njano Uyole)	51.73	51.73	51.73	33.92	-18.56
	(0.134)	(0.135)	(0.134)	(0.430)	(0.856)
	[0.146]	[0.146]	[0.146]	[0.626]	[0.887]
Apron Star Sachet (1=Sachet)	546.13***	546.13***	546.13***	534.52***	471.42***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
	[0.000]	[0.000]	[0.000]	[0.027]	[0.001]
Apron Star Pre-Treatment (1 = Pre-Treated)	762.27***	762.27***	762.27***	745.37***	532.66***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)
	[0.000]	[0.000]	[0.000]	[0.015]	[0.002]
DPTP x Seed variety				35.98	44.22
				(0.581)	(0.486)
				[0.594]	[0.541]
DPTP x Apron Star sachet				23.46	34.14
				(0.824)	(0.732)
				[0.833]	[0.762]
DPTP x Apron Star Pre-Treatment				34.14	65.06
				(0.679)	(0.407)
				[0.661]	[0.445]
Mbozi district		205.99	185.32	-42.73	198.03
		(0.196)	(0.178)	(0.811)	(0.239)
		[0.286]	[0.0236]	[0.808]	[0.310]
Mbozi x Seed variety					-9.95
					(0.874)
					[0.881]
Mbozi x Apron Star sachet					-25.59
					(0.807)
					[0.834]
Mbozi x Apron Star Pre-Treatment					2.38
					(0.974)
					[0.974]
Education (Respondent)		32.76	51.01**	24.11	33.19
		(0.234)	(0.019)	(0.489)	(0.199)
		[0.242]	[0.030]	[0.514]	[0.179]
Education x Seed variety					8.23
					(0.243)
					[0.264]
Education x Apron Star sachet					10.89
					(0.382)
					[0.415]
Education x Apron Star Pre-Treatment					30.24**
					(0.048)
					[0.090]
Bid order 2	-77.18	-10.03	-60.88	-69.99	-60.88
	(0.791)	(0.974)	(0.826)	(0.802)	(0.826)
	[0.786]	[0.980]	[0.813]	[0.787]	[0.813]
Bid order 3	-587.73**	-556.17**	-589.41**	-585.27**	-589.41**
	(0.026)	(0.044)	(0.017)	(0.017)	(0.017)
	[0.035]	[0.059]	[0.019]	[0.020]	[0.019]

Table 2A.3 (cont'd)

	(1)	(2)	(3)	(4)	(5)
Bid order 4	-430.20	-365.38	-386.95	-403.94	-386.95
	(0.125)	(0.171)	(0.159)	(0.131)	(0.160)
	[0.138]	[0.190]	[0.166]	[0.146]	[0.166]
Bid order 5	-632.52**	-590.40**	-607.54**	-610.05**	-607.54**
	(0.028)	(0.043)	(0.026)	(0.031)	(0.026)
	[0.028]	[0.052]	[0.024]	[0.027]	[0.024]
Bid order 6	-914.21***	-848.48**	-862.84***	-823.71***	-862.84***
	(0.007)	(0.012)	(0.007)	(0.009)	(0.007)
	[0.009]	[0.013]	[0.009]	[0.009]	[0.009]
DPTP x Mbozi				426.23*	
				(0.078)	
				[0.107]	
DPTP x Education				37.40	
				(0.348)	
				[0.379]	
Age		-8.54			
		(0.176)			
		[0.226]			
Gender (1 = Male)		64.09			
		(0.561)			
		[0.597]			
Served on Village Council (1 = Yes)		153.16			
		(0.417)			
		[0.483]			
Total land owned (acres)		-1.17			
		(0.971)			
		[0.974]			
Share of titled land (%)		-0.17			
		(0.967)			
		[0.969]			
PPI Poverty Likelihood (%)		0.77			
		(0.881)			
		[0.879]			
Asset Index Score		-36.21			
		(0.499)			
		[0.543]			
Constant	2,911.57***	2,805.56***	2,437.67***	2,757.12***	2,574.25***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	[0.000]	[0.000]	[0.000]	[0.002]	[0.001]
Observations	1,746	1,746	1,746	1,746	1,746
R-squared	0.128	0.161	0.151	0.160	0.152

Notes: OLS regressions. Dependent variable is the farmer bid for a given product in Tanzanian Shillings (Tsh). Standard errors are clustered at the village level (12 clusters). Robust p-values in parentheses. In square brackets we report wild cluster p-values (Wu 1986) generated using boottest command in Stata 14 (Roodman 2015).
*** p<0.01, ** p<0.05, * p<0.1 for the wild cluster p-values respectively.

Table 2A.4: Effects of treatment on knowledge excluding the control group

VARIABLES	(1) Uyole 96	(2) Njano Uyole	(3) Apron Star
Demonstration Plot (DP)	0.16	0.05	0.90***
	(0.724)	(0.818)	(0.000)
	[0.747]	[0.838]	[0.000]
Demonstration Plot + Trial Packs (DPTP)	0.31	0.34	0.54***
	(0.367)	(0.222)	(0.000)
	[0.430]	[0.252]	[0.008]
Age	-0.00	-0.01*	-0.00
	(0.258)	(0.069)	(0.722)
	[0.303]	[0.086]	[0.752]
Education (Respondent)	0.05**	-0.02	-0.04
	(0.025)	(0.263)	(0.349)
	[0.032]	[0.326]	[0.324]
Gender (1 = Male)	-0.25	-0.06	0.07
	(0.108)	(0.666)	(0.769)
	[0.109]	[0.664]	[0.770]
Served on Village Council (1 = Yes)	0.15	0.32	0.50*
	(0.385)	(0.123)	(0.051)
	[0.381]	[0.129]	[0.066]
Total land owned (acres)	-0.01	0.01	-0.01
	(0.421)	(0.524)	(0.822)
	[0.454]	[0.502]	[0.854]
Share of titled land (%)	0.01*	0.01*	-0.01**
	(0.023)	(0.055)	(0.042)
	[0.058]	[0.080]	[0.020]
PPI Poverty Likelihood (%)	0.00	0.00	-0.00
	(0.669)	(0.535)	(0.705)
	[0.676]	[0.580]	[0.698]
Asset Index Score	0.11***	0.03	0.05
	(0.005)	(0.324)	(0.343)
	[0.004]	[0.332]	[0.404]
Mbozi district	0.24	0.24	-0.35***
	(0.465)	(0.248)	(0.004)
	[0.455]	[0.253]	[0.005]
Constant	-0.26	-0.11	-1.52**
	(0.564)	(0.804)	(0.023)
	[0.565]	[0.811]	[0.052]
Observations	435	435	435
Hypothesis Testing			
H_0: DP = DPTP	[0.725]	[0.247]	[0.009]

Notes: Probit regressions. Dependent variable is farmer awareness of a given technology. Standard errors are clustered at the village level (12 clusters). Robust p-values in parentheses. In square brackets we report wild cluster p-values (Wu 1986) generated using boottest command in Stata 14 (Roodman 2015).
*** p<0.01, ** p<0.05, * p<0.1 for the wild cluster p-values.

APPENDIX 2B: Bid order effects

Previously presented results control for the random order in which farmers were asked to bid on improved bean products in this experiment. We used six possible bid orderings that vary the order of products by both bean variety and the Apron Star technology implemented (Table 2B.1). For example, in bid orders one and two farmers first place bids on untreated seed, then on untreated seed paired with a sachet, followed by pre-treated seed. We expect farmer WTP for an improved bean seed technology to be independent of any experimental design variables, including bid order.

Table 2B.1: Product bid orders

Bid order	Untreated Uyole 96	Untreated Njano Uyole	Untreated Uyole 96 w/ Apron Star Sachet	Untreated Njano Uyole w/ Apron Star Sachet	Pre-treated Uyole 96	Pre-treated Njano Uyole
1	1	2	3	4	5	6
2	2	1	4	3	6	5
3	3	4	1	2	5	6
4	4	3	2	1	6	5
5	6	5	4	3	2	1
6	5	6	3	4	1	2

However, we find evidence that the bid order used did, in fact, have a significant impact on farmer WTP for improved bean technologies. Table 2.4 includes coefficients for the effect of experimental order on bidding behavior. When compared to the first auction order, we find evidence that bid orderings three through six significantly decrease farmer WTP when compared to the first ordering. For example, in column 1, exposure to bid order five reduces WTP by an average of 577.4 Tsh. This reduction is large in magnitude, equivalent to a 19.6% decline in the estimated valuation for untreated Uyole 96. Our pattern of estimated coefficients on auction order provides some evidence that bidding on products including Apron Star early-on can decrease average WTP. However, when comparing the joint significance of *all* bid orders, we fail to reject the null hypothesis of joint equality to zero (F-test p-value = 0.152). Overall, this finding suggests that order effects may not be a major issue in estimating farmer WTP for these improved bean technologies

Our findings of some significant individual order effects are similar to those observed by Morawetz, De Groote, and Kimenju (2011) who implemented a BDM mechanism with Kenyan consumers of white and yellow maize flour. During their test round, the authors find evidence of a significant order effect that is not present during the main round. They argue that order effects occur when individuals do not understand the mechanism and focus on the need for implementing a practice round prior to bidding. Our study incorporated a practice round using a bar of soap, a common household item, helping to mitigate concerns about participant understanding.

A second potential explanation is that significant order effects might indicate learning about new products by bidding (Lusk and Shogren 2007).[33] Our findings appear to be more in line with a learning explanation, where individuals learn about the products as they place bids, and more importantly, set a reference point for their other submitted valuations. A farmer who uses an initial bid for untreated seed as an anchor is likely to have, on average, higher bids for seeds with some level of Apron Star treatment than producers who bid first on packages including Apron Star.

We further explore the effect of auction bid order on farmer WTP for improved bean technologies by comparing farmer bids within a specific auction bid order.[34] Table 2B.2 presents the results of running separate regressions that restrict the sample to farmers that faced the same bid order. Here, column number corresponds to the bid order subsample presented in Table 2B.1. Across all six bid orders and consistent with our main results, we still find little evidence that our

[33] Participant fatigue is also a potential source of significant order effects (Lusk and Shogren 2007). Our auction process kept participant time to a minimum to mitigate bidding fatigue. We did however, conduct a one-hour survey prior to household participation in the BDM which could be a driver of significant order effects.
[34] Note, however, that restricting the sample of interest by auction order reduces statistical power.

two extension treatments significantly affect farmer WTP for bean technology bundles. However, for bid order 3 (column 3) we find evidence of extension treatment effects: both the demonstration plot and demonstration plot with trial packs significantly reduce farmer WTP by 440.6 Tsh and 402.6 Tsh respectively ($p<0.1$). After accounting for the small number of clusters however, these results are only weakly significant. The same could be said for the significant negative effect of the demonstration plot with trial pack treatment we find in bid order 6 (column 6).

Auction bid order also appears to affect farmer WTP for product attributes. Focusing first on seed variety, we reject the null hypothesis that coefficient estimates in our six models are equal ($p<0.05$). Farmers submitting bids under the first auction order (column 1) are willing to pay an average of 103.8 Tsh less for Njano Uyole relative to Uyole 96 ($p<0.1$). However, farmers exposed to the second (column 2), fourth (column 4), and fifth (column 5) auction orders are willing to pay significantly more, on average, for Njano Uyole than Uyole 96 (110.2 Tsh ($p<0.1$), 93.1 Tsh ($p<0.1$), and 102.6 Tsh ($p<0.05$), respectively). These findings are consistent with results presented in Table 2.6, where we find evidence of a small positive increase in WTP for Njano Uyole on average compared to Uyole 96.

We find less heterogeneity due to bid order in farmers' WTP for bean seed bundled with a sachet of Apron Star. First, consistent with estimates presented in Table 2.6, we find that farmers are willing to pay an average price premium of 463.5 Tsh ($p<0.01$) to 690.0 Tsh ($p<0.01$) for improved bean seed bundled with an Apron Star sachet. However, we do reject the null hypothesis of equality of coefficient estimates across the six models ($p<0.01$). Instead, we find that farmer WTP for seed bundled with an Apron Star sachet is highest in auction orders three and four (columns 3 and 4 respectively), where farmers submit bids for these technology packages first.

Our findings for the effect of bid order are similar for WTP for pre-treated seed. Consistent with earlier results (Table 2.6), we find strong evidence that farmers have a higher WTP for pre-treated seed compared to untreated across all six models. However, there is a large range of estimated effects from 510.0 Tsh (column 2) to 1,043.7 Tsh (column 4) depending on the bid order farmers were subjected to. We again reject the null hypothesis that the coefficients across the six models are equal ($p>0.1$) and find a similar pattern as with seed with a sachet of Apron Star for self-treatment where auction orders three and four result in higher WTP estimates for pre-treated seed.

Differences in WTP due to the random order in which farmers submit bids might influence our finding that farmers are willing to pay significantly more for pre-treated seed than for seed bundled with a sachet of Apron Star for self-treatment. Across auction orders three, four, and six, we find that farmers are willing to pay a statistically significant price premium for pre-treated seed over self-treated seed ($p<0.05$). However, this difference is not statistically significant among farmers submitting bids using orders one, two, or five. Thus, our overall price premium for seed treatment found in Table 2.6 appears to be driven largely by the higher WTP for pre-treated seed induced by auction orders three and four.

Table 2B.2: Effect of auction order on farmer WTP for bean technologies (OLS estimates)

	Order 1	Order 2	Order 3	Order 4	Order 5	Order 6
Demonstration Plot (DP)	-250.43	600.09	-440.58*	291.41	-231.96	-457.69
	(0.520)	(0.077)	(0.053)	(0.490)	(0.533)	(0.153)
	[0.576]	[0.140]	[0.084]	[0.566]	[0.581]	[0.190]
Demonstration Plot + Trial Packs (DPTP)	243.39	-33.13	-402.61*	260.13	-303.14	-735.75*
	(0.489)	(0.898)	(0.085)	(0.546)	(0.358)	(0.040)
	[0.522]	[0.906]	[0.095]	[0.611]	[0.434]	[0.056]
Seed variety (1=Njano Uyole)	-103.83*	110.19*	2.22	93.12*	102.55**	5.21
	(0.062)	(0.065)	(0.963)	(0.049)	(0.034)	(0.910)
	[0.065]	[0.054]	[0.955]	[0.050]	[0.031]	[0.908[
Apron Star Sachet (1=Sachet)	556.56***	463.51***	690.00***	684.92***	556.18***	521.09***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	[0.000]	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]
Apron Star Treatment (1 = Treated)	661.48***	510.00***	1,028.67***	1,043.65***	645.88***	703.12***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	[0.000]	[0.001]	[0.000]	[0.000]	[0.001]	[0.000]
Mbozi district	474.60	536.36**	86.10	335.71	596.75**	216.98
	(0.123)	(0.027)	(0.615)	(0.333)	(0.011)	(0.475)
	[0.122]	[0.030]	[0.646]	[0.406]	[0.017]	[0.508]
Education level (Respondent)	78.02**	67.25*	-49.60	68.26	51.36	-19.87
	(0.043)	(0.035)	(0.187)	(0.101)	(0.186)	(0.663)
	[0.039]	[0.068]	[0.322]	[0.106]	[0.203]	[0.688]
Constant	2,246.23***	1,903.42***	2,869.15***	1,413.60**	1,873.73***	2,576.39***
	(0.000)	(0.000)	(0.000)	(0.011)	(0.001)	(0.000)
	[0.005]	[0.000]	[0.000]	[0.011]	[0.001]	[0.000]
Observations	366	522	450	378	510	384
R-squared	0.159	0.162	0.152	0.183	0.142	0.144
Hypothesis Testing			**p-value**			
H0: Apron Star Sachet = Apron Star Treatment	[0.269]	[0.641]	[0.001]	[0.002]	[0.316]	[0.018]

Notes: OLS regressions. Dependent variable is the farmer bid for a given product in Tanzanian Shillings (Tsh). Standard errors are clustered at the village level (18 clusters). Robust p-values in parentheses. In square brackets we report wild cluster p-values (Wu 1986) generated using boottest command in Stata 14 (Roodman 2015). *** p<0.01, ** p<0.05, * p<0.1 for the wild cluster p-values respectively.

REFERENCES

REFERENCES

Abadie, A., S. Athey, G. Imbens, and J. Wooldridge. 2017. "When Should You Adjust Standard Errors for Clustering?" *arXiv:1710.02926 [econ, math, stat]*. Available at: http://arxiv.org/abs/1710.02926 [Accessed May 11, 2018].

Aker, J.C. 2011. "Dial 'A' for agriculture: a review of information and communication technologies for agricultural extension in developing countries." *Agricultural Economics* 42(6):631–647.

Anderson, J., and L. Crowder. 2000. "The present and future of public sector extension in Africa: contracting out or contracting in?" *Public Administration and Development* 20(5):373–384.

Anderson, J.R. 2008. "Agricultural Advisory Services." *World Bank Background paper*:44.

Anderson, J.R., and G. Feder. 2007. "Agricultural Extension." In *Handbook of Agricultural Economics*. pp. 2343–2378.

Anon. 2017. "Apron Star 42WS Seed Care." Available at: https://www.syngenta.co.ke/print/551.

Bandiera, O., and I. Rasul. 2006. "Social Networks and Technology Adoption in Northern Mozambique." *The Economic Journal* 116(514):869–902.

Barrett, C.B., L. Christiaensen, M. Sheahan, and A. Shimeles. 2017. "On the Structural Transformation of Rural Africa." *Journal of African Economies* 26(suppl_1):i11–i35.

Beaman, L., A. BenYishay, J. Magruder, and A. Mobarak. 2015. "Can Network Theory-based Targeting Increase Technology Adoption?" *Working Paper*. Available at: http://www.ferdi.fr/sites/www.ferdi.fr/files/pictures/beaman_et_al_-can_network_theory-based_targeting_increase_technology_adoption.pdf [Accessed May 17, 2018].

Becker, G., M.H. DeGroot, and J. Marschak. 1964. "Measuring Utility by a Single-Response Sequential Method." *Behavioral Science* 9(3):226–232.

Benin, S., E. Nkonya, G. Okecho, J. Randriamamonjy, E. Kato, G. Lubade, and M. Kyotalimye. 2011. "Returns to spending on agricultural extension: the case of the National Agricultural Advisory Services (NAADS) program of Uganda." *Agricultural Economics* 42(2):249–267.

Bensch, G., M. Grimm, and J. Peters. 2015. "Why do households forego high returns from technology adoption? Evidence from improved cooking stoves in Burkina Faso." *Journal of Economic Behavior & Organization* 116:187–205.

BenYishay, A., and A. Mobarak. 2015. "Social Learning and Incentives for Experimentation and Communication." Available at:

http://faculty.som.yale.edu/mushfiqmobarak/papers/MalawiAg.pdf [Accessed May 17, 2018].

Blackie, M., and K. Albright. 2005. "Lesson learning study of the Farm Inputs Promotions (FIPS) project in Kenya (with a special emphasis on public-private partnerships for input provision and possibilities for regional upscaling)." :19.

Bruhn, M., and D. McKenzie. 2009. "In Pursuit of Balance: Randomization in Practice in Development Field Experiments." *American Economic Journal: Applied Economics* 1(4):200–232.

Cameron, A.C., J.B. Gelbach, and D.L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *The Review of Economics and Statistics* 90(3):414–27.

Colin Cameron, A., and D.L. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources* 50(2):317–372.

Conley, T., and C. Udry. 2010. "Learning about a New Technology: Pineapple in Ghana." *The American Economic Review* 100(1):35–69.

Davis, K., E. Nkonya, E. Kato, D.A. Mekonnen, M. Odendo, R. Miiro, and J. Nkuba. 2012. "Impact of Farmer Field Schools on Agricultural Productivity and Poverty in East Africa." *World Development* 40(2):402–413.

De Groote, H., S.C. Kimenju, and U.B. Morawetz. 2011. "Estimating consumer willingness to pay for food quality with experimental auctions: the case of yellow versus fortified maize meal in Kenya." *Agricultural Economics* 42(1):1–16.

Dupas, P. 2014. "Short-Run Subsidies And Long-Run Adoption Of New Health Products: Evidence From A Field Experiment." *Econometrica* 82(1):197–228.

Emerick, K., and M.H. Dar. 2017. "Enhancing the diffusion of information about agricultural technology." *Unpublished Manuscript*:48.

Esarey, J., and A. Menger. 2018. "Practical and Effective Approaches to Dealing With Clustered Data." *Political Science Research and Methods*:1–19.

Foster, A., and M. Rosenzweig. 1995. "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture." *Journal of Political Economy* 103(6):1176–1209.

Gibson, I.S.A. 1953. "Crown rot, a seedling disease of groundnuts caused by Aspergillus niger: II. An anomalous effect of organo-mercurial seed dressings." *Transactions of the British Mycological Society* 36(4):324–334.

Goeb, J. 2017. *Information, Knowledge, and Demand for Substitute Health Inputs: Experimental Evidence of Pesticide Use in Zambia*. Michigan State University.

Hillocks, R.J., C.S. Madata, R. Chirwa, E.M. Minja, and S. Msolla. 2006. "Phaseolus Bean Improvement in Tanzania, 1959–2005." *Euphytica* 150(1–2):215–231.

Hoffmann, V., C.B. Barrett, and D.R. Just. 2009. "Do Free Goods Stick to Poor Households? Experimental Evidence on Insecticide Treated Bednets." *World Development* 37(3):607–617.

Kondylis, F., V. Mueller, and J. Zhu. 2017. "Seeing is believing? Evidence from an extension network experiment." *Journal of Development Economics* 125:1–20.

Lusk, J., and J.F. Shogren. 2007. *Experimental auctions: methods and applications in economic and marketing research*. Cambridge ; New York: Cambridge University Press.

Lybbert, T.J., N. Magnan, A.K. Bhargava, K. Gulati, and D.J. Spielman. 2013. "Farmers' Heterogeneous Valuation of Laser Land Leveling in Eastern Uttar Pradesh: An Experimental Auction to Inform Segmentation and Subsidy Strategies." *American Journal of Agricultural Economics* 95(2):339–345.

Magnan, N., D.J. Spielman, T.J. Lybbert, and K. Gulati. 2015. "Leveling with friends: Social networks and Indian farmers' demand for a technology with heterogeneous benefits." *Journal of Development Economics* 116:223–251.

Masters, W.A., and D. Sanogo. 2002. "Welfare Gains from Quality Certification of Infant Foods: Results from a Market Experiment in Mali." *American Journal of Agricultural Economics* 84(4):974–989.

Melkani, A., and N.M. Mason. 2018. "Tanzania Southern Highlands Farm Input Promotions-Africa Village-Based Agricultural Advisors Baseline Survey Report." *Technical Report*:57.

Morawetz, U.B., H.D. Groote, and S.C. Kimenju. 2011. "Improving the Use of Experimental Auctions in Africa: Theory and Evidence." *Journal of Agricultural and Resource Economics* 36(2). Available at: http://www.jstor.org/stable/23243080.

Moser, C.M., and C.B. Barrett. 2003. "The disappointing adoption dynamics of a yield-increasing, low external-input technology: the case of SRI in Madagascar." *Agricultural Systems* 76(3):1085–1100.

Muhamba, G., L. Amos, D. Protas, P. Mbogo, and S. Nchimbi-Msoll. 2013. "Marker Assisted Selection for Common Bean Diseases Improvements in Tanzania: Prospects and Future Needs." In S. B. Andersen, ed. *Plant Breeding from Laboratories to Fields*. InTech. Available at: http://www.intechopen.com/books/plant-breeding-from-laboratories-to-fields/marker-assisted-selection-for-common-bean-diseases-improvements-in-tanzania-prospects-and-future-nee [Accessed May 12, 2018].

Munshi, K. 2004. “Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution.” *Journal of Development Economics* 73(1):185–213.

Nakano, Y., T.W. Tsusaka, T. Aida, and V.O. Pede. 2018. “Is farmer-to-farmer extension effective? The impact of training on technology adoption and rice farming productivity in Tanzania.” *World Development* 105:336–351.

Pan, Y., S.C. Smith, and M. Sulaiman. 2018. “Agricultural Extension and Technology Adoption for Food Security: Evidence from Uganda.” *American Journal of Agricultural Economics*. Available at: https://academic.oup.com/ajae/advance-article/doi/10.1093/ajae/aay012/4982756 [Accessed April 27, 2018].

Roodman, D. 2015. “BOOTTEST: Stata module to provide fast execution of the wild bootstrap with null imposed.” *Statistical Software Components*. Available at: https://ideas.repec.org/c/boc/bocode/s458121.html [Accessed June 6, 2018].

Schreiner, M. 2016. “Tanzania 2011 Progress out of Poverty Index® (PPI®): Design Memo.” :231.

Shenkalwa, E.M., M.E.T. Mmbaga, A. Kalala, and E. Shenkalwa. 2013. “Performance Of Improved Bean Varieties In Kasulu And Kibondo Districts Of Kigoma Region, Tanzania.” *African Journal of Food, Agriculture, Nutrition and Development* 13(3):16.

Suri, T. 2011. “Selection and Comparative Advantage in Technology Adoption.” *Econometrica* 79(1):159–209.

Tjernström, E. 2015. “Signals, Similarity and Seeds: Social Learning in the Presence of Imperfect Information and Heterogeneity∗.” *Unpublished Manuscript*:42.

Trutmann, P., K.B. Paul, and D. Cishabayo. 1992. “Seed treatments increase yield of farmer varietal field bean mixtures in the central African highlands through multiple disease and beanfly control.” 11:7.

Waldman, K.B., J.M. Kerr, and K.B. Isaacs. 2014. “Combining participatory crop trials and experimental auctions to estimate farmer preferences for improved common bean in Rwanda.” *Food Policy* 46:183–192.

Webb, M. 2014. “Reworking Wild Bootstrap Based Inference for Clustered Errors.” *Queen’s Economics Department Working Paper*:1–21.

Whitfield, S., J.L. Dixon, B.P. Mulenga, and H. Ngoma. 2015. “Conceptualising farming systems for agricultural development research: Cases from Eastern and Southern Africa.” *Agricultural Systems* 133:54–62.

Wu, C.F.J. 1986. "Jackknife, Bootstrap, and Other Resampling Methods in Regression Analysis." *Annals of Statistics* 14:1261–1350.

3. AGRICULTURAL INNOVATION AND TECHNOLOGY ADOPTION: INCENTIVES FOR DELAY RECONSIDERED

3.1 Introduction

Public and private investments in research and development have been key drivers of U.S. agricultural productivity growth, realizing total factor productivity (TFP) gains of approximately 1.5 percent per year since 1948 (Wang et al. 2015). Attempting to keep pace with population growth and projected increases in demand for food and fiber, policymakers have pushed new funding initiatives for agricultural research and development with the goal of raising innovation rates and giving producers more technology options (de Janvry and Dethier 1985; Alston et al. 2011; Pardey et al. 2015). Despite increased numbers of innovations being made available to farmers, a lingering puzzle has been the lower than expected rates of adoption and diffusion of ostensibly profitable agricultural technologies (e.g. Moser and Barrett 2003; Baerenklau and Knapp 2007; Duflo, Kremer and Robinson 2011; Suri 2011). Among potential explanations for this phenomenon are the characteristics of the innovation process itself. High transactions costs, steep learning curves, and uncertain returns associated with farmer adoption of new technologies make many new technologies a risky investment for farmers. A rapidly changing choice set driven by technological innovation might influence individual preferences over the adoption of currently available and future technologies. This is especially important when farmers face highly variable rates of innovation in agricultural technologies (e.g., Michigan saw 18 improved soybean varieties released in 2015 compared to 30 in 2016) (Tjernström 2015; Wang et al. 2015). The goal of this investigation is to further develop the empirical links between the characteristics of innovation processes and technology adoption decisions. In particular, this research seeks to answer the question: how does the rate of technological innovation influence the probability of technology adoption as well as the number of new technologies individuals take up?

We focus on how two key uncertainties may drive farmers to delay the adoption of profitable new agricultural technologies. The first is uncertainty about the rate of innovation. New technologies are the culmination of a long process of research and development activities including product testing, certification, and production. Unexpected holdups or knowledge spillovers can delay or advance the innovation process at multiple levels, suggesting that the arrival of innovations on the marketplace is a stochastic process (Grenadier and Weiss 1997). Because investment in agricultural technologies is often characterized by some irreversibility, uncertainty surrounding the arrival of new innovations can cost farmers in two ways. One cost is the direct cost of making an investment prematurely and facing a superior technology in the near future (Farzin et al. 1998). Switching technologies can be costly, especially when secondary markets for selling a technology are thin. A second issue is the opportunity cost of delaying an investment in a new agricultural technology. Conditional on an innovation being profitable, each period or season that a farmer does not invest in a new superior technology while waiting on a future innovation represents some lost revenue that could be gained from adoption (conditional on the new innovation being profitable).

The second key uncertainty analyzed here relates to the value of a new technology to individual producers. While farmers may expect new innovations, if adopted, to increase on-farm returns or lower costs, there is no guarantee of the magnitude (or even sign) of these benefits once a technology is incorporated into the production process. The benefits of adoption could depend upon farm characteristics (e.g. soil quality, elevation), household characteristics (e.g. farmer education, experience), or even the use of complementary agricultural technologies. When the future returns to new technologies are unknown, farmers face a risk of delaying adoption of current technologies only to realize a disappointing new technology (e.g. lower return or higher variance) in a future period.

To gain intuition on how these two uncertainties may work together, consider four illustrative cases in the context of a high-yielding variety (HYV) seed technology. In the first and most important case, suppose that a farmer believes that there is a high probability of a more profitable HYV seed technology being released in the next period and that the expected net returns to this new seed will be large given the conditions faced by the farmer. The farmer will have a strong incentive to delay adoption of currently available technologies with the expectation of large future gains if the anticipated superior technology materializes. Because innovation is a self-reinforcing process, new innovation could then strengthen incentives for farmers to face a similar dilemma in period $t + 1$. In the second case, suppose that the opposite is true, and farmers believe there is a low probability of a superior HYV seed being released in the next period and, if it materializes, the expected increase in future profits would be negligible. In this case, expected future gains from delay will increase by a small amount and are unlikely to outweigh the benefits of adopting an existing seed technology in the current period. Finally, there are two indeterminate cases where either the probability of superior HYV innovations is high and the expected net benefits are low, or the probability is low and the expected net benefits are high. In both of these instances, the effect on future gains will depend on the relative magnitudes of the adoption cost, the probability of innovation arrival, and the expected return of the future innovation. We would expect that for these two cases the likelihood of delaying adoption would be small given the discounting of future earnings.

We design a lab experiment to identify the effect of the rate of innovation on individual technology adoption behavior. Subjects are tasked with choosing production technologies to maximize total returns over a period of 26 rounds. The main treatment of interest is the rate at which new technologies are generated and made available for use. In the high (low) innovation rate environment, there is an 80% (20%) chance that a new technology is generated at the beginning of any given period and can be acquired at a fixed cost. Returns to a given new technology evolve according to a stochastic process where subjects can expect a new innovation to, on average, have a higher mean return and a lower variance of returns. Combined, these features of the choice environment allow us to analyze the effects of the innovation rate on the probability, number, and timing of technology adoption decisions.

Our results suggest that increasing the rate of innovation may provide some incentives for individuals to pass on the adoption of current new technology offerings in favor of waiting and adopting new technologies representing larger improvements in expected returns (over and above existing technologies) as soon as they arrive. We show that increasing the rate of innovation increases the probability that individuals adopt new technologies upon arrival while decreasing the probability of adopting a new innovation with delay or adopting a previously used technology. Furthermore, we find that an increase in the rate of innovation significantly reduces the total number of new technologies individuals adopt. Together, these results suggest that higher rates of innovation may push individuals into an attrition strategy, where they delay adoption decisions until a sufficiently improved technology is made available.

This research makes several contributions to the existing literature. First, we experimentally measure observed individual adoption behavior under a change in the rate of technological innovation. While there have been several studies that numerically simulate changes in adoption patterns (Grenadier and Weiss 1997; Farzin, Huisman and Kort 1998; Doraszelski 2001; Doraszelski 2004), we place subjects in an incentivized decision environment. Second, our approach has an advantage over empirical studies of technology timing that use observational data (Carletto, de Janvry and Sadoulet 1999; Carletto et al. 2010; Huh and Kim 2008; Kim and Srinivasan 2009) because we eliminate any effects of marketing strategies for existing and future technologies that might be confounded with the uncertain arrival of new innovations. Third, unlike previous simulation efforts, we allow for the partial reversibility of technological investment by implementing tiered fixed adoption costs. Individuals face a high fixed cost to adopt a technology they have never used before, and a lower fixed cost to adopt a previously used technology. This is a useful contribution because we focus on the sunk costs of adoption rather than technology lock-in. Fourth, we test the effect of individual risk aversion on adoption behavior. Rather than assuming risk neutrality as in the simulation-based studies, we measure individual risk preferences using a lottery choice module prior to conducting the main experiment, then control for individual risk aversion in our analysis of individual adoption behavior. Some empirical analyses characterize a new technology as inherently more risky than traditional versions (Carletto et al. 1999; Carletto et al. 2010) but use non-adoption as a proxy for risk aversion rather than using a direct measure of risk aversion as we do here.

3.2 Literature review

This article builds on several lines of research related to the rate of technological innovation: technology adoption and diffusion, strategic entry into a technology, the option value of delay, and empirical studies of technology adoption timing. We highlight key studies in each of these areas.

3.2.1 Agricultural technology adoption and diffusion

Beginning with Griliches (1957), agricultural economists have long been interested in understanding the adoption behavior of producers and the diffusion of agricultural technologies. Farmer technology adoption decisions have been shown to depend on heterogeneous characteristics across farmers and farms including farm size, land tenure, soil quality, and credit access (Feder, Just and Zilberman 1985; Feder and Umali 1993; Foster and Rosenzweig 2010). A particular focus in this literature has been on why producers fail to adopt ostensibly profitable technologies. Major findings identify heterogeneity in net returns (Suri 2011), uninsured risk (Karlan et al. 2014), time-inconsistent preferences (Duflo et al. 2011), lack of access to credit (Feder, Just, and Zilberman 1985; Giné and Yang 2009), and imprecise learning from observation and experimentation (Foster and Rosenzweig 1995; Munshi 2004). On the supply side of new technologies, Emerick et al. (2016) show that suitability of a primary technology to a production environment, can significantly crowd-in the use of complementary improved inputs and management practices. In our study, we use technologies that generate the same level of returns for all individuals and provide a known distribution of returns to control for many issues related to learning and heterogeneity in returns due to producer characteristics.

Fewer studies have focused on the dynamics of technology adoption decisions over time, and most studies treat adoption as a one-time decision (McWilliams and Zilberman 1996). There are a few important exceptions. Social learning studies suggest there may be incentives to strategically delay adoption decisions to increase profitability in the presence of learning from others (Foster and Rosenzweig 1995; Munshi 2004; Conley and Udry 2010). Specifically, when the optimal level of use of a technology is uncertain, a producer may want to gather more information about its use by observing others before investing in an innovation. Once the investment is made, the additional knowledge from the earlier delay may lead to higher profitability than if the user relied only on own experimentation (Foster and Rosenzweig 1995). While we do not address any of the social learning mechanisms in this experiment, we do allow individuals to make multiple adoption decisions over time.

The availability of new innovations in the agricultural sector are linked to public and private investments in research and development (R&D). Motivating large investments are high estimated returns to agricultural R&D spending, with conservative estimates placing mean returns around 10% (Alston et al. 2000; Alston et al. 2010; Alston et al. 2011; Hurley, Rao and Pardey 2014). More recently however, there has been a great deal of concern surrounding declining agricultural productivity growth in the U.S. and abroad that has been attributed to declining investment in the agricultural sector (Alston et al. 2010; Ball, Schimmelpfennig, and Wang 2013; Pardey et al. 2015). Cai, Golub and Hertel (2017) call for substantial acceleration of R&D spending through 2050 based on simulations of the productivity increases that will be needed to account for future supply and demand uncertainty. Increasingly, agricultural R&D spending is shifting towards innovation streams in developing countries as well as towards private sector investments (Pardey et al. 2015).

R&D spending translates into an increased number of technologies for farmers in a variety of ways. First, public funding through national governments, international CGIAR centers, and land-grant universities often generates new technologies, such as in the form of productivity enhancing genetics or improved management practices, that can be broadly disseminated and adapted for local use in many locations. Second, private investment often results in branded products and services marketed to agricultural producers. Farm machinery, patented plant genetics (e.g. Roundup Ready), agricultural chemicals, and precision agriculture equipment and affiliated software are examples of this output. Increasingly, the private sector accounts for a larger share of total agricultural R&D spending (44%), but this is still relatively low when compared to other sectors of the economy (Pardey et al. 2015). Finally, there is a growing literature (Biggs 1990; Spielman 2005; Neef and Neubert 2011; Kilelu, Klerkx and Leeuwis 2013) suggesting farmers may be a source of innovation in and of themselves. Experimentation with and adaptation of existing technologies can create effective new technologies. This could be something as simple as a new management practice (e.g. fertilizer micro dosing) or modification of existing equipment for a new purpose. Our study builds on this literature by analyzing individual behavior in response to a sharp increase in the arrival of new agricultural innovations similar to those that might be driven by increases in R&D funding.

3.2.2 Strategic adoption under rivalry

The timing of technology adoption decisions also has important implications for the welfare of end-users. Cochrane (1958) describes a treadmill pattern where early farm-technology adopters earn larger returns due to increased productivity or lower costs. As technologies diffuse, production increases and output prices fall; the profitability gap between early and late adopters decreases until all farmers are forced to adopt to stay in the market and non-adopters exit.

The treadmill dynamic introduces a strategic dimension into firm adoption behavior where there may be distinct first- or second-mover advantages to investing in a new innovation (see the review by Hoppe 2002). Using a game-theoretic approach, Reinganum (1981) finds that firms competing in a duopolistic setting in the presence of declining adoption costs will adopt a productivity enhancing technology at different dates in the presence of a first-mover production advantage. Huisman and Kort (2004) extend this model to show that when the gains to preemption are small, there is a symmetric strategy involving late adoption of the technology. They add the arrival of a second new production technology to the duopoly setting. Irreversible investments paired with the uncertain arrival of a second innovation delays the adoption decision of both firms. Optimal firm strategy depends on the probability of the new technology's arrival. At low arrival probabilities, the preemption game dominates. At higher arrival probabilities the game becomes an attrition game where one or both firms will wait for the new innovation to arrive and ignore the current technology (Huisman and Kort 2004). This experiment focuses on similar uncertain arrival times, but in a decision theoretic framework. We allow individuals to adopt new technologies without considering strategic interaction in a manner more akin to perfect competition. Further extensions of the strategic interaction framework has focused on informational spillovers (Hoppe 2000), information heterogeneity (Yoon 2009), and discontinuous payoff structures (Smirnov and Wait 2015) that can generate significant incentives to be the second mover in an adoption decision.

Harou, Walker, and Barrett (2017) argue that while a first-mover advantage holds for an unambiguously and permanently superior technology, this may not be true for agricultural production where profitability is highly dependent on consumer demand. Comparison of non-adopters to late adopters of new pineapple technologies in Ghana reveal small differences in profitability and welfare gains in the presence of market shocks.[35] We would expect the characteristics of the technology under consideration (e.g. mean and variance of returns) as well as the market environment for agricultural output to influence the relative benefits of adopting or upgrading a new agricultural technology. For now, however, we assume perfectly competitive output markets where all individuals receive the same price for output.

[35] Harou, Walker, and Barrett (2017) measure welfare differences using an asset index and find less than a 0.1 standard deviation difference between late-adopters who experience a market shock and non-adopters.

3.2.3 Analyzing dynamic adoption decisions

Option value models (Dixit and Pindyck 1994) provide an analytical framework that is well suited to understanding the timing of technological investments by explicitly modeling investment under uncertainty and irreversibility. These models suggest that there is an extra hurdle, in addition to positive net present value (NPV), that new technologies must overcome to trigger an adoption decision (McDonald and Siegel 1986). Specifically, the NPV criteria predicts that an individual will adopt a technology when the expected present value of returns exceeds the present value of any investment costs. This behavior is rarely observed in the field (Caswell and Zilberman 1985; Duflo et al. 2011; Suri 2011; Moser and Barrett 2003) where individuals can take advantage of an option to delay the investment until a later date.

Real options models have been applied to a variety of technology adoption decisions. Purvis et al. (1995) show that the option value of waiting to invest in free-stall cattle housing dominated the value of investing, potentially slowing down the diffusion of new cattle technologies. Here, option value means the time value in the decision to delay a costly investment in free-stall cattle housing. The producer has the right, but not the obligation, to make the investment similar to a financial option. Similar findings have been reported for hybrid seed varieties (Dong and Saha 1998), irrigation systems (Carey and Zilberman 2002), and site-specific farming technologies (Isik and Khanna 2003). Building on these previous efforts, Baerenklau and Knapp (2007) incorporate technology age, reversible investments, variable inputs, and stochastic prices into a numerical simulation and show that: (i) relaxing irreversibility assumptions promotes technology adoption; and (ii) the delaying impact of price uncertainty declines with technology age.

Most closely related to our study is a robust literature focusing on the timing of investment decisions in the presence of stochastic technology returns and arrival times. Early work focused solely on uncertain returns to a new technology whereby an individual can delay a costly investment while receiving independent signals about an innovation's profitability (Jensen 1982). Once a critical upper (lower) level of confidence is reached the technology will be adopted (rejected) (McCardle 1985). Costly experimentation and information gathering increases the complexity of the managerial decision (Jensen 1988; Hoppe 2002) and is similar to the learning process about optimal technology use proposed by Foster and Rosenzweig (1995).

Balcer and Lippman (1984) were the first to directly model uncertain technology arrival times for a technological improvement. Using a discrete-time semi-Markov process for innovation, they find two behavioral regularities for technology adoption. First, when the speed of technological innovations increases, firms may avoid switching costs and postpone an adoption decision. Second, as time passes with no new innovations arriving, it may become profitable to purchase an existing superior technology that was previously passed over. Similarly, Weiss (1994) uses a discrete time model to argue that individual expectations of larger technological improvements in the future might incentivize the delay of technological adoptions. When expected gains are large enough, the firm will adopt the future improvement as soon as it is innovated, ending the search process.

Extensions of these models have focused directly on the option value of delaying a technology investment decision when returns evolve over time. Grenadier and Weiss (1997) consider a continuous time model where a firm makes adoption decisions over a current technology and an uncertain future innovation. They find the probability of adopting the current technology is decreasing in the speed of innovation. Faster innovation encourages firms to bypass current technologies or delay the purchase until the future technology is realized. Similarly, increases in the expected returns to a future innovation increase the probability of either waiting for the innovation to arrive or purchasing each innovation sequentially. Farzin et al. (1998) present an environment where the investment decision is fully irreversible (e.g. the firm makes a single choice to lock-in a technology) and there is uncertainty over the arrival time and degree of technological improvement of a new innovation. Technological innovation is assumed to always increase payoffs, disallowing the case of technological regression. They find that a new technology will be adopted immediately whenever its value exceeds a threshold that is significantly higher than the NPV rule, even in the case allowing a finite number of technology switches (Farzin et al. 1998; Doraszelski 2001).[36] If instead the value of a new innovation is below the threshold, the innovation will be bypassed (a leapfrog strategy) in favor of a future technology. Building upon these previous studies, we allow for the partial reversibility of technological investment in the decision environment by implementing tiered fixed adoption costs. Individuals face a high fixed cost to adopt a technology they have never used before, and a lower fixed cost to adopt a previously used technology. This structure allows individuals to more freely switch technologies, a major shortcoming of previous simulation efforts.

Doraszelski (2004) extend the real options framework to distinguish between technological innovations and technological improvements. An innovation "triggers a myriad of small improvements that enhance the efficiency of the basic technology" (Doraszelski 2004, Pg. 1462). Improvements are small incremental increases in the mean returns to a technology that are distinguished from innovations in their time dependency. Specifically, the arrival time of improvements is dependent on the time elapsed since the last innovation, with improvements to a

[36] Immediate adoption of the new technology occurs due to discounting, where the firm could always improve on the overall payoff by adopting the new superior innovation earlier.

current technology becoming less frequent over time. Numerical simulation reveals that a firm might not wait on the next breakthrough technology but can maximize returns by waiting for a current technology to become sufficiently advanced (even as the probability of an improvement declines over time). We do not explicitly leverage this feature of time dependency in this article, instead choosing to view the arrival of new innovations as independent and identically distributed draws from a Bernoulli distribution. Overall, while these simulations inform our expectations over individual behavior, we test the probability and timing of observed technology adoption decisions in an experimental environment.

3.2.4 Empirical studies of innovation speed and adoption timing

A related literature focuses on the timing of agricultural technology decisions using hazard functions to estimate the determinants of time to adoption (or disadoption) (Carletto et al. 1999; Burton, Rigby and Young 2003; Abdulai and Huffman 2005; Carletto et al. 2010). For example, Abdulai and Huffman (2005) estimate survival models for the length of time required for Tanzanian farmers to adopt crossbred-cow technology as a function of household and farm characteristics. They find that time to adoption is decreasing in education, age, herd size, and nonfarm income but increasing in the technology price and distance to market. One of the drawbacks to these studies remains their focus on adoption as a single-spell issue where, once adopted, the household is out of the market for new technologies. In reality, however, we often see producers moving in and out of agricultural technology use, especially for those with low fixed costs (Chen and Myers 2017). We control for several demographic characteristics in this model to better understand individual heterogeneity in technology adoption decisions.

For many consumer electronics, there is significant evidence that consumer beliefs about the rate of innovation matter for individual adoption decisions. Focusing on consumer cell phone purchases in Korea, Huh and Kim (2008) find that more intensive use of a technology's advanced or innovative features leads to a higher intention to purchase the next generation of the technology. Kim and Srinivasan (2009) were among the first to focus explicitly on the determinants of timing in technology upgrade decisions for personal digital assistants (PDAs). They find strong evidence that individuals who believe PDA technology will improve more quickly in the future tend to delay upgrading their current technologies. Mukherji, Rajagopalan, and Tanniru (2006) simulate the adoption problem for a deterministic innovation process where a firm can choose to increase output by upgrading software every year. They find that optimal time to upgrade is increasing in the technology cost and decreasing in the productivity benefits (opportunity cost).

3.3 Conceptual framework

We consider an individual in period t is presented with k_t available technologies that can be used in a generic production process. Returns to a technology are given by R_t^k which has mean of μ_k and a variance of σ_k^2.[37] We assume that returns in any given period are drawn from a uniform distribution where:

$$R_t^k \sim Uniform(\mu_k - \frac{\sigma_k^2}{2}, \mu_k + \frac{\sigma_k^2}{2}) \tag{3}$$

A uniform distribution is appropriate because a producer might have a reasonable expectation of the extreme values possible with technology use but may have no expectations over the exact return that will be achieved. We also assume that all other production inputs are fixed, and the

[37] In a discrete uniform distribution the variance is directly proportional to the range.

individual decision over which technology to use in the production process is the only explicit choice variable.

Production takes place in an environment of technological innovation which implies that the number of technologies available on the marketplace, k_t, is increasing over time. Innovation evolves according to a Bernoulli process:

$$k_{t+1} = \begin{cases} k_t + 1 \; with \; probability \; q \\ k_t \; with \; probability \; 1 - q \end{cases} \tag{2}$$

where an individual will observe a new innovation in any period with probability q. Because the number of available technologies is the sum of independent Bernoulli trials, the expected number of new innovations during any time period can be represented via a Poisson distribution as the limiting case. Furthermore, inter-arrival times of new technologies would be exponentially distributed and an individual's expectations over the probability of arrival should be time invariant (Farzin et al. 1998; Doraszelski 2001).

Conditional on an innovation being generated, we allow the mean of returns to evolve over time following a random walk with drift:

$$\mu_{k+1} = \mu_k + \varepsilon, \;\; \varepsilon \sim N(\mu, \sigma^2) \tag{3}$$

where μ_{k+1} is the mean return of the new innovation, μ_k is the mean return of the most recent innovation, and ε is the normally distributed drift parameter. We allow for positive drift in the innovation mean by setting $\mu > 0$ so that in expectation, individuals expect future innovations to have higher returns than current innovations. We allow for uncertainty in expected gains to an innovation by setting the variance of the drift parameter, σ^2. Contrary to previous studies (e.g. Grenadier and Weiss 1997; Farzin et al. 1998; Doraszelski 2001; Doraszelski 2004) we do not restrict future innovations to be positive changes in mean technological efficiency. The suitability of a new innovation for an individual's production process is not guaranteed as there can be significant inefficiencies generated by the interaction of a new innovation with the fixed characteristics or complementary technologies used in a production process. For example, a new variety of hybrid seed released on the market may be poorly suited to the agro-ecological conditions of a farmer's land (e.g. soil quality) and, if used, might result in a decline in expected yield.

We implement a similar random walk process with drift to describe the variance (range) of returns associated with a new innovation:

$$\sigma_{k+1}^2 = \sigma_k^2 + \eta, \;\; \eta \sim N(\lambda, \tau^2) \quad (4)$$

where σ_{k+1}^2 is the variance of the new innovation, σ_k^2 is the variance of the most recent technology to arrive, and η is the normally distributed drift parameter. Here we allow for negative drift in innovation variance by setting $\lambda < 0$. On average, individuals will expect a smaller range of returns, or more precision, from new innovations. We set the variance of the new innovation τ_k^2 at a level such that an increase in the range of returns (a loss of precision) is possible.

In our model, there are costs associated with switching from one technology to another. There are no restrictions on switching behavior, in contrast to the assumption of full irreversibility used in other work that limits the individual or firm to either a single or sequential investment (Grenadier and Weiss 1997; Farzin et al. 1998; Doraszelski 2001). The first type of cost, s_t^N, is the switching cost associated with adopting a new, never before used technology. Searching for the new technology, installation, developing technology management practices, and learning about optimal use would be included in this term. The second cost, s_t^O, is the switching cost associated with adopting a previously used technology. These costs can include re-optimization of production, re-learning production management processes, etc. We also assume there are no other costs associated with the continued use of a technology (i.e. once adopted, a technology will continue to provide returns at no maintenance cost).[38] We order switching costs such that:

$$s_t^N > s_t^O > 0 \tag{5}$$

The reduced cost of adopting a previously used technology is to reflect some participant experience or learning that would likely make switching back less costly than switching to a brand-new technology. Switching costs, s_t, are captured by a state variable and are conditional on what type of switch has occurred:

$$s_t = \begin{cases} s_t^O \; if d_t^k \in (d_1^k, \dots, d_{t-2}^k) \\ s_t^N if d_t^k \notin \left(d_1^k, \dots, d_{t-1}^k\right) \\ 0 \; if \; d_t^k = \; d_{t-1}^k \end{cases} \tag{6}$$

We have a single choice variable, d_t^k, which denotes the choice over the k_t available technologies:

$$d_t^k = \begin{cases} 1 \; if \; k \; is \; the \; chosen \; action \\ 0 \; otherwise \end{cases} \tag{7}$$

We assume individuals are risk averse and chooses production technologies to maximize the expected utility of the discounted sum of returns over an infinite horizon. State variables include discrete number of available technologies k_t, returns to the chosen technology R_t^k, switching costs s_t, and conditional on a new innovation, the mean μ_k and variance σ_k^2 of the new technology. The problem we seek to solve is:

$$\max_{d_t^k} EU\left[\left(\sum_{t=0}^{\infty} \beta^t \sum_{k=1}^{k_t} R_t^k d_t^k - s_t\right) | k_{t,}\mu_k, \sigma_k^2\right] \tag{8}$$

We can solve the problem using dynamic programming. The value function when comparing discrete technology choices is:

$$V_t(d_{t-1}) = \max\left\{V_t^1(d_{t-1}^1), \dots, V_t^k\left(d_{t-1}^k\right)\right\} \tag{9}$$

[38] This assumption could be relaxed by depreciating returns over time when a technology is used in production (or adding a term to capture increasing maintenance costs).

where $V_t^k(d_{t-1})$ is the present value if technology k is chosen at t, conditional on the state space. The conditional value function for technology k is specified as:

$$V_t^k(d_{t-1}) = \mathrm{EU}\left(R_t^k - s_t\right) + \beta EU(V_{t+1}(d_t)) \tag{10}$$

where $V_{t+1}(d_t)$ is the discounted value of future profits from choosing technology k today, and assuming optimal technology choices are made in the future.

Due to the stochastic arrival of innovations which drive an increase in the state space over time, an analytic solution to the optimal policy rule is intractable. Comparison of the conditional value functions over different scenarios can be instructive for understanding how we might expect individuals to behave in different innovation environments. We focus on a case where a new innovation arrives. Suppose an individual used technology i in period $t-1$. At the beginning of period t, suppose a new technology j arrives and is available for adoption. An individual will choose to switch and adopt technology j if the following holds:

$$EU\left(R_t^j\right) - EU\left(R_t^i\right) > s_t^N + \beta\left[EU\left(V_{t+1}\left(d_t^i\right)\right) - EU\left(V_{t+1}\left(d_t^j\right)\right)\ \right] \tag{11}$$

The left-hand side (LHS) illustrates the difference in expected returns from switching to technology j. The right-hand size (RHS) of equation (11) decomposes the costs of switching. The first term represents the switching cost associated with adopting a new (to the individual) technology. The second term on the RHS is the discounted expected future returns from continuing to use technology i in the current period. Through this second term, the individual also takes into account the expectation that a new technology with higher average returns also arrives in a future period.

In the presence of switching costs, an increase in the rate of innovation paired with technological improvement (expected increases in mean technological efficiency) would decrease

$EU\left(V_{t+1}\left(d_t^j\right)\right)$. This decline would occur because there would be a relatively high probability of another superior innovation arriving quickly and the individual incurring the switching cost again (i.e. incurring additional sunk costs). Additionally, adopting technology j in the current period might make adoption of a future incremental innovation irrational, resulting in additional opportunity cost. This would increase the value of delaying the decision to a later period and would make adoption of technology j less likely.[39]

Technological regression, the arrival of a new technology with lower expected mean returns than existing technologies, would incentivize the continued use of the current technology by making the LHS of equation (11) negative and the RHS larger (and positive). Furthermore, the realization of a new innovation with a lower mean could even motivate a switch to a previously existing unadopted technology to the extent that it increases the expected time until a new technology with a sufficient improvement in expected returns arrives. A faster rate of innovation increases the probability of experiencing a technological regression, and would be more likely to incentivize adoption of a preexisting market technology.

Based on this model for how innovation might incentivize delay, we test two different hypotheses: **(H1)** An increase in the rate of innovation will increase the probability that an individual delays the adoption of a new technology in favor of waiting for a better new innovation to arrive. Here, delayed adoption suggests that an individual will be more likely to forego adopting a technology in a given round in favor of waiting to adopt a more significant technological improvement if and when it arrives. Relative to a low innovation rate environment, individuals exposed to faster innovation will be more likely continue to use the same technology during any

[39] However, if the improvements in technological efficiency (LHS) are sufficiently large to outweigh the switching costs and the discounted future returns associated with delaying adoption for one period and waiting, the individual will immediately adopt.

given period. At the same time, we would expect a higher rate of innovation to make individuals less likely to adopt an innovation as soon as it arrives precisely because they are more willing to wait on future improvements. **(H2)** An increase in the rate of innovation will result in a lower number of technologies adopted over time for all measures and types of technology adoption. This hypothesis is driven by the expectation that over the same technology set, subjects in the high innovation rate treatment are likely to forego a larger number of incremental improvements than those in the low innovation rate treatment because the expected arrival time of an improvement is relatively sooner and the opportunity cost of delay is relatively lower in the high innovation rate environment. Passing up current technologies for future innovations will result in fewer technologies adopted in the high innovation rate group.

3.4 Experimental design and data

3.4.1 Experimental design parameters

The experiment was implemented using oTree software (Chen, Schonger and Wickens 2016) and run online with students at Michigan State University (MSU). Students were recruited from a sample of undergraduates enrolled at MSU using ORSEE (Greiner 2015). Treatments were conducted in three sessions, with 40 students invited to participate in each session (20 per treatment). In total, there were 112 participants who completed all portions of the experiment, with 58 in a high innovation treatment and 54 in a low innovation treatment. All sessions in this experiment included two stages. Stage 1 involved a standard lottery choice experiment following Holt and Laury (2002) to elicit a measure of risk preference.[40] In Stage 2, subjects participated in

[40].See Appendix 3B Table 3B.1 for an overview of the decisions and possible point earnings in the lottery choice experiment.

a 26 round technology choice environment with the between-subject treatment conditions (high innovation versus low innovation) imposed. At the end of the experiment, subjects completed a short demographic survey and were paid their earnings from their decisions in the form of a gift card to an online retailer. Full experimental instructions for each treatment are provided in Appendix 3A.

We implement two treatments over the rate of technological innovation. In the high innovation treatment, there is an 80% chance that a new technology is invented at the beginning of a round. In the low innovation environment, the probability of a new technology is set at 20%. Subjects are provided with a menu of technology options that can be used in the production environment. Once generated, a new technology is immediately added to a subject's choice set and remains available for use throughout the duration of the experiment. Table 3.1 presents the innovation patterns generated using these parameters and faced by subjects in each treatment.

Excluding the baseline technology immediately available during the first round, subjects in the high innovation treatment observe 19 new technologies during the course of the 26 round experiment compared to 10 in the low innovation treatment. Subjects are informed about the rate of innovation they will face at the beginning of the experiment, and the probability of observing a new technology in any period is independent of previous innovations.

Table 3.1: Technology generation patterns

Round	High Innovation (p=0.8)	Low Innovation (p = 0.2)
1	Baseline	Baseline
2	Technology A	Technology A
3	Technology B	Technology B
4	Technology C	
5		Technology C
6	Technology D	Technology D
7	Technology E	
8	Technology F	
9	Technology G	Technology E
10	Technology H	
11		
12		Technology F
13		
14	Technology I	
15	Technology J	
16		Technology G
17	Technology K	
18	Technology L	
19		Technology H
20	Technology M	
21	Technology N	Technology I
22	Technology O	Technology J
23	Technology P	
24	Technology Q	
25	Technology R	
26	Technology S	

Following equation (1), returns to a technology in each round are determined by a draw from a discrete uniform distribution. Subjects are presented with the mean and range of possible returns associated with each technology on the decision screen. We allow the mean and variance of returns to new innovations to evolve over time following equations (3) and (4) respectively. Mean returns follow a random walk process with drift, where the drift parameter, ε, is normally distributed with a mean of 2 points and a standard deviation of 8 points. On average, we can expect that new technologies will have higher mean returns than older technologies, but this is not guaranteed. The range of returns (variance) follows a similar evolution process; however, we set the drift parameter to be normally distributed with a mean of -2 points and a standard deviation of 8 points. We would then expect new technologies to have a smaller range of returns, on average, compared to existing technologies, reflecting more predictable outcomes when using new innovations. But, again, it is not guaranteed that every new technology will have a small range of returns. Full details on the mean and range of returns to each new technology can be found in Table 3.2. Note that we hold the stochastic process governing the evolution of returns to new innovations constant across the two treatments. A subject using Technology C in the high innovation treatment will face the same uniform distribution for returns as a subject using Technology C in the low innovation treatment. This ensures that we are only changing the rate of the innovation process between the two treatment groups.

Technology adoption in this experiment requires the payment of a fixed cost, making the adoption decision only partially reversible. Subjects seeking to adopt a new technology that they have not used in a previous round must pay an up-front cost of 50 points. Once paid, an individual can continue using the technology in future rounds at no cost. This cost is the same if the technology arrived in the current round or was added to the choice set previously. Individuals choosing to adopt a technology previously used in another period can switch back for a cost of 25 points. Subjects are provided with a starting endowment of 100 points to cover the cost of two changes to technologies not used previously. This is to prevent individuals who would like to adopt a new technology as soon as it becomes available from having to delay the decision until they have accumulated enough points to pay the fixed cost. The baseline technology used in both rounds is provided at no cost to participants.

Table 3.2: Technology point returns

Technology	Mean returns	Minimum	Maximum	Range of returns
Baseline	80	0	160	160
A	91	9	174	165
B	103	28	178	150
C	114	40	189	149
D	139	70	209	139
E	154	92	217	125
F	159	101	217	116
G	153	92	214	122
H	156	94	218	124
I	165	98	233	135
J	163	100	226	126
K	172	115	230	115
L	172	114	230	116
M	171	117	225	108
N	186	128	245	117
O	176	118	234	116
P	175	114	236	122
Q	184	129	240	111
R	180	131	230	99
S	186	140	233	93

At the end of each round, the players viewed a summary screen with the following information: (i) the technology they used during the round; (ii) the returns generated by the technology; (iii) whether or not they switched production technologies, and the level of the switching cost incurred; and (iv) total cumulative earnings after the current round.[41] Point earnings were totaled across all 26 rounds (and the lottery choice experiment), and converted to dollars at a rate of $1.00 per 500 points. At the end of the experiment, subjects were presented with a

[41] Additionally, after every fourth round subjects were asked about their expectations for seeing a new innovation during the next period and how they expected the mean and variance of returns to compare to current technology offerings.

performance summary describing their total earnings, then they completed the short survey collecting demographic information.

3.4.2 Summary Statistics

Table 3.3 presents summary statistics across our pooled sample and for the high versus low innovation treatments for several experimental outcomes as well as for the demographic variables collected during the experiment. We also test for differences in mean values between the high and low treatment groups using nonparametric Mann-Whitney tests.

Table 3.3: Summary statistics by treatment

	Full Sample (N=112)			High Innovation (N=58)			Low Innovation (N=54)				
	mean	min	max	mean	min	max	mean	min	max	Difference	p-value
Experimental outcomes											
Experimental earnings (points)	3,674	2,174	4,456	3,785	2,761	4,291	3,554	2,174	4,456	231***	0.001
Number of technology adoptions	7.116	0	21	7.172	1	21	7.056	0	21	0.116	0.329
Number of technologies adopted immediately	5.527	1	19	6.052	1	19	4.963	1	11	1.089*	0.090
Number of technologies adopted with delay	1.384	0	8	1.362	0	8	1.407	0	8	-0.045	0.626
Number of previously used technologies adopted	1.196	0	12	0.759	0	11	1.667	0	12	-0.908	0.312
Average number of rounds between adoption	5.566	1.4	21	5.568	1.48	21	5.564	1.4	12.81	-0.004	0.701
Relative risk aversion coefficient	0.273	-0.950	1.370	0.209	-0.950	1.370	0.343	-0.950	1.370	-0.134	0.275
Demographic Characteristics											
Age	21.36	19	28	21.34	19	26	21.37	19	28	-0.030	0.936
Gender (1=Male)	0.357	0	1	0.293	0	1	0.426	0	1	-0.133	0.145
Student status	4.045	2	6	4.103	2	6	3.981	2	6	0.122	0.460
In-state student (1=Yes)	0.938	0	1	0.931	0	1	0.944	0	1	-0.013	0.771
Survey responses											
Ever delayed technology purchase to wait for a new model? (1=Yes)	0.741	0	1	0.707	0	1	0.778	0	1	-0.071	0.394
How likely are you to adopt a new cell phone technology as soon as it arrives?	2.312	1	5	2.397	1	5	2.222	1	4	0.175	0.323
Do you agree that new technologies will be released faster in the future?	4.027	1	5	4	1	5	4.056	2	5	-0.056	0.560

Note: [a] Student status is coded as: 1=Freshman, 2=Sophomore, 3=Junior, 4=Senior, 5=Graduate student, 6=Alumni. [b] Responses are coded as: 1=Extremely unlikely, 2=Unlikely, 3=Neither likely nor unlikely, 4=Likely, 5=Extremely likely. [c] Responses are coded as: 1=Strongly disagree, 2=Disagree, 3=Neither agree nor disagree, 4=Agree, 5=Strongly agree. Significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1, respectively. P-values are from the results of a nonparametric Mann-Whitney U test. Results from a two-tailed t-test are similar.

Participants are of similar age, student status, and whether or not a student is from the state of Michigan. Our sample is more heavily weighted towards females (64.3% overall), however we observe no statistically significant differences in the proportion of men and women assigned to the high and low innovation treatments ($p>0.1$). Additionally, we find no statistically significant differences in subject experiences with consumer technologies and beliefs over the rate of innovation between the two treatment groups ($p>0.1$). Survey questions reveal that 74.1% of subjects have delayed a technology purchase to wait on a new release; subjects report being slightly unlikely to adopt a new cell phone technology upon arrival; and, on average, participants believe that the rate of innovation is likely to increase in the future.

Relative risk aversion coefficients are calculated from the Stage 1 lottery choice experiment, where a range of relative risk aversion values are calculated for the number of risk and safe lottery choices an individual makes. Following Holt and Laury (2002), we use the midpoint of the range of relative risk aversion in our measure. Note that higher values of the relative risk aversion coefficient imply that an individual is more risk averse. The difference in average risk aversion exhibited between the high and low innovation treatments using our measure is not statistically different from zero ($p>0.1$).

Subjects in the high innovation treatment earned significantly higher returns on average (3,785 points) during the course of the experiment compared to the low innovation treatment (3,554 points) ($p<0.01$). Additionally, we find some evidence that individuals in the high innovation treatment adopt 1.09 more technologies as soon as they are invented ($p<0.1$). Both of these results are unexpected because subjects in the high innovation environment could adopt higher return innovations earlier and observed additional technological improvements in later rounds that the low innovation treatment did not see.

To check for balance across the treatment groups on participant observables, we regress the treatment assignment (1 = High, 0=Low) on the full set of demographic variables collected in the survey. We then run an F test for the null hypothesis that observed covariates are all equal to zero against the alternative hypothesis that at least one of the coefficients is statistically different from zero (Bruhn and McKenzie 2009). Failure to reject the null hypothesis is evidence that the treatment groups are balanced. We fail to reject the null hypothesis that the treatment groups are balanced based on demographic characteristics in Table 3.4, column 1. In column 2, we add the survey responses on participants' experiences with consumer technology and again fail to reject the null of balanced treatment groups. Because the survey was conducted after subjects were exposed to the experimental treatment, it is possible that individual expectations over future technological innovation were influenced by the treatment status. Therefore, in the results that follow, we use only the demographic variables (and not the consumer technology-related questions) as controls.

Table 3.4: Regression-based balance test based on observable participant characteristics

	(1) High vs. Low Innovation	(2) High vs. Low Innovation
Age	-0.062	-0.062
	(0.045)	(0.045)
Gender (1=Male)	-0.118	-0.101
	(0.101)	(0.103)
Sophomore	-0.031	-0.012
	(0.204)	(0.210)
Junior	-0.132	-0.089
	(0.298)	(0.304)
Senior	-0.227	-0.181
	(0.223)	(0.228)
Graduate student	-0.147	-0.117
	(0.177)	(0.184)
In-state student (1=Yes)	0.229	0.243
	(0.194)	(0.196)
Ever delayed tech purchase to wait for a new model? (1=Yes)		-0.152
		(0.128)
How likely are you to adopt a new cell phone technology as soon as it arrives?		0.044
		(0.053)
Do you agree that new technologies will be released faster in the future?		-0.010
		(0.052)
Constant	2.004*	2.000*
	(1.051)	(1.083)
Observations	112	112
R-squared	0.059	0.073
F-Test		
H_0: Joint Orthogonality (p-values)	0.491	0.557

Note: OLS regressions. The dependent variable is equal to one if a subject is in the high innovation group and zero if in the low innovation group. Alumni is the excluded reference category for student status. Standard errors in parentheses. Significance is denoted b: *** p<0.01, ** p<0.05, * p<0.1, respectively. F-test is used to test the null hypothesis that the coefficients of all demographic variables are jointly equal to zero against the alternative that at least one coefficient is statistically different from zero. Failure to reject the null hypothesis is evidence that the treatment groups are balanced on observable covariates.

3.5 Results

Each of the 112 subjects made technology adoption decisions in every round of the experiment. Given 26 rounds in the technology adoption portion of the experiment, this results in a dataset of 2,912 adoption decisions pooled across the two treatments. In the following sections, we will explore the effect of treatment on two main outcomes of interest: the probability of making specific unordered adoption decisions within a given round and the number of technologies adopted over multiple rounds.[42]

3.5.1 Probability of technology adoption

We first consider the effect of changing the rate of innovation on the probability that a subject adopts a new technology in a given round. We consider the four types of adoption scenarios that can occur within a round: *No Change*, *Adopt Now, Adopt Delay,* and *Adopt Used*. *No Change* occurs when a subject chooses to continue using a current technology in a given round – there is no technology switch and no fee incurred. *Adopt Now* is when a subject adopts a new technology in the round when it is first made available. *Adopt Delay* is when a subject adopts a new technology in any round after the round in which it is first made available. Finally, *Adopt Used* is when a subject re-adopts a technology that was previously used (and abandoned) in earlier rounds. Together, these four measures are mutually exclusive and cover all of the possible adoption decisions of a participant in a given round.

To estimate the effect of increasing the rate of innovation on participant adoption decisions, we estimate the following multinomial logit model:

[42] We also explore the effect of treatment on the average duration of technology use and the average duration of delay. See Appendix 3C for details.

$$P(Adopt_{it} = j) = \frac{\exp\,(\boldsymbol{Z_{it}\beta_j})}{\sum_j \exp\,(\boldsymbol{Z_{it}\beta_j})}\ for\ j = 0,1,\dots,J \tag{12}$$

$$\begin{aligned} \boldsymbol{Z_{it}\beta_j} = \beta_{0j} + \beta_{1j}High_i + \beta_{2j}Risk_i + \beta_{3j}Good_{i,t-1} + \beta_{4j}Innov_{i,t-1} \\ + \beta_{5j}Switch_{i,t-1} + \boldsymbol{Round_t\varphi_j} \end{aligned} \tag{13}$$

where the dependent variable of interest, $Adopt_{it}$ is one of the four unordered outcome choices subject i can make in round t (*No Change*, *Adopt Now, Adopt Delay, Adopt Used*). To identify the parameters in the model, *No Change* is specified as the reference category, and coefficient estimates are relative to this outcome choice. $\boldsymbol{Z_{it}}$ is the vector of independent variables, expanded in equation (13). Our main treatment variable of interest is $High_i$ and is equal to one if the subject is in the high innovation rate treatment and zero for the low innovation rate treatment. $Risk_i$ controls for the level of relative risk aversion of a participant as estimated from the results of the Stage 1 lottery choice experiment. We also control for individual experiences in the previous round that might affect adoption in current round: $Good_{i,t-1}$ is equal to one if the stochastic return from the technology used in the last period was greater than or equal to the expected return, and zero otherwise. $Innov_{i,t-1}$ is a binary variable indicating if a new technology arrived in the previous round and $Switch_{i,t-1}$ is a binary indicator for if the individual changed technologies in the previous round. $\boldsymbol{Round_t}$ is a vector of dummy variables controlling for the round.[43]

We present coefficient estimates for the effect of increasing the rate of innovation on the probability of choosing a specific adoption outcome in a given round in Table 3.5 and we present the corresponding average partial effects (APEs) on adoption outcome probabilities in Table 3.6.

[43] Subject demographic controls were excluded from the multinomial logit specification due to convergence issues.

Focusing first on the coefficient estimates in Table 3.5, Columns 1-3 present estimates using adoption data on all technologies presented to subjects in both innovation treatments. Relative to the reference category of *No Change*, we find that an increase in the rate of innovation decreases the probability that an individual adopts a previously used technology ($p<0.1$), decreases the probability that an individual adopts a new technology with some measure of delay ($p<0.05$), and increases the probability that an individual adopts a new technology in the round it first becomes available ($p<0.05$). This pattern of results provides little support for **(H1)** where a high rate of innovation is expected to decrease the probability of adopting a new technology as soon as it becomes available. Instead, we find the opposite that a higher rate of innovation makes individuals more likely, on average, to adopt a new technology as soon as it arrives relative to continuing to use a current technology.

Table 3.5: Multinomial logit coefficient estimates for adoption behavior in a given round

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Sample	All technologies			Common technologies			First five technologies		
Dependent Variable	Adopt Used	Adopt Delay	Adopt New	Adopt Used	Adopt Delay	Adopt New	Adopt Used	Adopt Delay	Adopt New
High innovation treatment	-0.659*	-0.602**	0.366**	-0.482	-0.185	0.886***	-0.790**	-0.903**	0.350*
	(0.058)	(0.047)	(0.024)	(0.178)	(0.562)	(0.000)	(0.042)	(0.022)	(0.093)
Relative risk aversion coefficient	-0.531***	-0.271*	-0.428***	-0.516***	-0.231	-0.399***	-0.489***	-0.162	-0.135
	(0.000)	(0.063)	(0.000)	(0.001)	(0.142)	(0.000)	(0.008)	(0.540)	(0.409)
Good tech return (t-1)	-0.430**	-0.412**	-0.337***	-0.456**	-0.325*	-0.543***	-0.472**	-0.351	-0.575***
	(0.024)	(0.017)	(0.003)	(0.018)	(0.085)	(0.000)	(0.043)	(0.266)	(0.003)
Innovation arrival (t-1)	-0.082	0.181**	-0.044	-0.204**	-0.161	-0.801***	-0.108	0.246	-16.541
	(0.370)	(0.035)	(0.372)	(0.043)	(0.139)	(0.000)	(0.428)	(0.277)	(0.972)
Technology switch (t-1)	2.109***	-0.090	0.061	2.039***	-0.077	-0.096	1.686***	-0.864**	-0.381*
	(0.000)	(0.654)	(0.659)	(0.000)	(0.725)	(0.537)	(0.000)	(0.019)	(0.099)
Constant	-19.422	-19.653	0.204	-20.574	-20.886	0.783***	-23.111	-24.217	16.742
	(0.995)	(0.995)	(0.331)	(0.997)	(0.997)	(0.000)	(0.999)	(0.999)	(0.972)
Round controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,799	2,799	2,799	2,799	2,799	2,799	2,800	2,800	2,800

Note: Dependent variable is the adoption behavior of an individual in a given round. In each round the individual has 4 choices: use their current technology (*No Change*), adopt a previously used technology (*Adopt used*), adopt a new technology from a previous round (*Adopt delay*), or adopt an innovation that just arrived (*Adopt now*). We exclude *No Change* as the reference category. Columns 1-4 use data on all technologies in the experiment. Columns 5-8 use data on only common technologies between the two treatments (Baseline Technology through Technology J). Columns 9-12 use data on only the first five technologies. P-values in parentheses. Significance denoted by *** p<0.01, ** p<0.05, * p<0.1, respectively.

We test the robustness of this finding against two other specifications. In columns 4-6, we restrict the analysis to the first 11 technologies which were common between the two treatments (Baseline Technology through Technology J). The only difference between the two treatments is the speed at which these 11 technologies entered the subject's choice set. Here we find no statistically significant effects of treatment on the probability of adopting a previously used innovation or an innovation with delay relative to the reference category. However, we see a large, and statistically significant increase in the probability of adopting an innovation upon arrival in the high innovation rate treatment group ($p<0.01$). Much like the first set of results, this suggests that increasing the rate of innovation makes individuals more likely adopters of the newest innovation – potentially one they have been waiting one or more rounds to realize.[44] We also restrict the analysis to the arrival of the first five technologies in columns 7-9. This is an important distinction because, as Table 3.1 illustrates, the first five technologies (Baseline Technology through Technology D) arrive in a very similar pattern across the two treatments so the primary source of heterogeneity is due to individual expectations over the probability of a new innovation arriving. Here, our pattern of results for an increase in the rate of innovation mirrors that of using all technologies where we find that individuals decrease the probability of adopting a previous innovation ($p<0.05$) or a new technology with delay ($p<0.05$) and increase the probability of adopting a new technology upon arrival ($p<0.1$), all relative to the reference category of continuing to use the current technology.

[44] We also restrict the analysis to only those technologies that occur in the middle of the experiment when the realized rates of innovation diverge between the two treatment groups and technology arrival in the low innovation treatment group slows down. The pattern and magnitude of these results is indistinguishable from our estimates using the common set of technologies in Table 3.5 columns 4-6.

Table 3.6 Multinomial logit APEs for the effect of treatment on the probability of technology adoption behavior

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Sample	All technologies				Common technologies				First five technologies			
Dependent Variable	No Change	Adopt Used	Adopt Delay	Adopt New	No Change	Adopt Used	Adopt Delay	Adopt New	No Change	Adopt Used	Adopt Delay	Adopt New
High innovation treatment	0.009	-0.028**	-0.033*	0.051***	-0.050**	-0.023*	-0.012	0.085***	0.018	-0.021	-0.014	0.017
	(0.730)	(0.045)	(0.054)	(0.006)	(0.027)	(0.081)	(0.366)	(0.000)	(.)	(.)	(.)	(.)
Relative risk aversion	0.072***	-0.018***	-0.010	-0.044***	0.056***	-0.018***	-0.006	-0.031***	0.019	-0.013	-0.002	-0.004
	(0.000)	(0.003)	(0.192)	(0.000)	(0.000)	(0.003)	(0.327)	(0.001)	(.)	(.)	(.)	(.)
Good tech return (t-1)	0.065***	-0.014*	-0.018**	-0.034**	0.069***	-0.015*	-0.010	-0.044***	0.036	-0.012	-0.004	-0.020
	(0.000)	(0.067)	(0.045)	(0.010)	(0.000)	(0.060)	(0.216)	(0.000)	(.)	(.)	(.)	(.)
Innovation arrival (t-1)	-0.000	-0.004	0.010**	-0.006	0.075***	-0.003	-0.002	-0.070***	0.569	0.029	0.042	-0.639
	(0.964)	(0.347)	(0.026)	(0.311)	(0.000)	(0.409)	(0.656)	(0.000)	(.)	(.)	(.)	(.)
Technology switch (t-1)	-0.070	0.087	-0.010	-0.007	-0.056	0.0834	-0.007	-0.020	-0.019	0.048	-0.013	-0.016
	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)
Observations	2,799	2,799	2,799	2,799	2,799	2,799	2,799	2,799	2,800	2,800	2,800	2,800

Note: Dependent variable is the adoption behavior of an individual in a given round. In each round the individual has 4 choices: use their current technology (*No Change*), adopt a previously used technology (*Adopt used*), adopt a new technology from a previous round (*Adopt delay*), or adopt an innovation that just arrived (*Adopt now*). Columns 1-4 use data on all technologies in the experiment. Columns 5-8 use data on only common technologies between the two treatments (Baseline Technology through Technology J). Columns 9-12 use data on only the first five technologies. P-values in parentheses. Significance denoted by *** p<0.01, ** p<0.05, * p<0.1, respectively.

Calculating the APE of treatment on the probability of choosing one of the four adoption states can provide a direct test of **(H1)** where we would expect to observe an increase in the probability of choosing to continue using a current technology (*No Change*) and a decrease in the probability of immediate adoption (*Adopt Now*) in the high innovation treatment. Columns 1-4 in Table 3.6 present the APEs of treatment on the relative probabilities of choosing any given category. Because the four outcomes are mutually exclusive within a given round, the marginal effects of treatment necessarily sum to zero because an increase in the probability of choosing one outcome will decrease the probability of choosing one or more of the other outcome alternatives. Using all technologies, an individual in the high innovation treatment is, on average, 2.8 percentage points less likely to adopt a previously used technology ($p<0.05$), 3.3 percentage points less likely to adopt a new innovation with delay ($p<0.1$), and 5.1 percentage points more likely to adopt a new innovation in the round it first becomes available, relative to an individual in the low innovation treatment. However, contrary to **(H1)** we do not find a statistically significant effect of treatment on the probability of continuing to use the current technology and find a large increase in the probability of immediate adoption. Restricting the analysis to common (first 11) technologies in columns 5-8, we find that individuals in the high innovation treatment are 5.0 percentage points less likely to continue using a current technology ($p<0.05$), 2.3 percentage points less likely to adopt a previously used technology ($p<0.05$), and 8.5 percentage points more likely to adopt a new innovation as soon as it arrives ($p<0.01$). Here, we see strong evidence to suggest individuals are less likely to continue using a current technology in a given round, and instead choose to upgrade to newly arriving innovations. We also look at APEs among the first five technologies and find some evidence to support the expected pattern of results in **(H1)** where individuals in the high innovation treatment are 1.8 percentage points more likely to continue using current technologies, 2.1 percentage points less likely to adopt a used technology, 1.4 percentage points less likely to adopt a new technology with delay, and 1.7 percentage points more likely to adopt a new innovation as soon as it arrives. However, due to the discrete nature of the data and the small number of rounds in which subjects adopt and use the first five technologies, we are unable to calculate standard errors surrounding the APEs of treatment in these cases.

We also find an interesting pattern of results for the effects of an individual's level of risk aversion and experience in previous rounds of the experiment. Focusing on the APEs for common technologies (Table 3.6 columns 5-8), we find that, on average, a one unit increase in an individual's coefficient of relative risk aversion increases the probability that an individual continues to use a current technology by 5.6 percentage points ($p<0.05$), decreases the probability of adopting a previously used technology by 1.8 percentage points ($p<0.05$), and decreases the probability of adopting new innovations when they arrive by 3.1 percentage points ($p<0.01$). While a one unit change in relative risk aversion would be a large shift (from extremely risk loving to extremely risk averse), it does suggest that risk averse individuals perceive uncertain innovation patterns differently and are generally less likely to adopt new innovations than less risk averse individuals. In the context of technology investments with a fixed cost, adopting a new technology is potentially risky because a significant improvement could arrive in the next period requiring payment of another fixed cost, should the individual choose to adopt that new technology. Similar to an increase in risk aversion, individuals who realized an above average return to a current technology in the previous period are 6.9 percentage points ($p<0.01$) more likely to continue using that technology, 1.5 percentage points ($p<0.01$) less likely to adopt a previous used technology, and 4.4 percentage points ($p<0.01$) less likely to adopt a new technology when it first arrives. Despite the independence of returns across periods, subjects are more likely to stick with a current technology after a good period suggesting that they may believe there is a higher probability of receiving a high return in future periods. Finally, observing a new innovation in a previous round increases the probability of continuing to use a current technology by 7.5 percentage points ($p<0.01$) and decreases the probability of adopting a new innovation immediately by 7.0 percentage points ($p<0.01$). Overall these findings suggest that individual characteristics like risk aversion and previous experiences with technologies (and realized innovations) can affect current adoption decisions.

3.5.2 Count of technologies adopted

We next consider how an increase in the rate of innovation affects the number of technologies adopted by individuals. Allowing for the possibility of heterogeneous effects, we focus on five different count variables that include: the number of unconditional technology adoptions allowing for any change (*Adopt*), the number of new innovations adopted (*Adopt New*), the number of new technologies adopted in the round they arrive (*Adopt Now*), the number of new technologies adopted with at least a one round delay (*Adopt Delay*), and the number of used technologies re-adopted (*Adopt Used*). We estimate the effects of treatment on the number of technologies adopted using Poisson regressions of the form:

$$E(Count_i \mid Treat_i, Risk_i, \boldsymbol{X}_i) = \exp(\beta_0 + \beta_1 High_i + \beta_2 Risk_i + \boldsymbol{X}_i\boldsymbol{\delta}) \tag{14}$$

where $Count_i$ is the number of technologies adopted by individual i during the experiment. $High_i$ is the binary treatment indicator, $Risk_i$ controls for the level of relative risk aversion of a participant as estimated in Stage 1, and $\boldsymbol{X}_i$ is a vector of participant demographic characteristics including subject age, gender, status as a Michigan resident, and student status at MSU.

Table 3.7: APEs of treatment on the number of technologies adopted (Poisson)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Sample	Common Technologies					First Five Technologies				
Dependent Variable	Adopt	Adopt New	Adopt Now	Adopt Delay	Adopt Used	Adopt	Adopt New	Adopt Now	Adopt Delay	Adopt Used
High innovation treatment	-2.508***	-1.260***	-0.674*	-0.544*	-1.309***	-1.135**	-0.260	-0.053	-0.209	-0.996**
	(0.001)	(0.003)	(0.088)	(0.059)	(0.010)	(0.014)	(0.196)	(0.808)	(0.134)	(0.011)
Relative risk aversion coefficient	-1.331*	-0.763*	-0.755*	-0.013	-0.491	-0.460	-0.074	-0.097	0.030	-0.354
	(0.093)	(0.072)	(0.054)	(0.965)	(0.350)	(0.314)	(0.666)	(0.619)	(0.820)	(0.349)
Age	0.308	0.079	-0.047	0.136	0.244	0.099	0.038	-0.059	0.089*	0.043
	(0.447)	(0.744)	(0.840)	(0.184)	(0.305)	(0.648)	(0.705)	(0.564)	(0.051)	(0.828)
Gender (1=Male)	-2.707***	-1.209***	-0.284	-0.884***	-1.558***	-1.304***	-0.349*	0.093	-0.424***	-1.002***
	(0.000)	(0.003)	(0.483)	(0.000)	(0.000)	(0.000)	(0.083)	(0.671)	(0.000)	(0.000)
Michigan resident (1=Yes)	-0.165	-0.099	-0.853	0.666***	0.241	0.350	-0.097	-0.632	0.454***	0.582**
	(0.901)	(0.890)	(0.280)	(0.007)	(0.633)	(0.559)	(0.809)	(0.126)	(0.000)	(0.015)
Sophomore	3.355	0.273	0.380	-0.101	4.816	0.821	-0.573	-0.684	0.260	2.897
	(0.284)	(0.858)	(0.800)	(0.850)	(0.208)	(0.564)	(0.314)	(0.273)	(0.496)	(0.304)
Junior	0.213	0.261	-0.127	0.444	0.043	-0.403	0.021	-0.158	0.205	-0.355
	(0.911)	(0.811)	(0.902)	(0.519)	(0.965)	(0.712)	(0.966)	(0.774)	(0.531)	(0.623)
Senior	-0.815	-0.495	-0.372	-0.113	-0.187	-0.985	-0.504	-0.430	-0.073	-0.313
	(0.603)	(0.525)	(0.632)	(0.835)	(0.817)	(0.290)	(0.157)	(0.339)	(0.787)	(0.633)
Graduate student	-0.230	0.040	0.359	-0.276	-0.140	-0.688	-0.578*	-0.326	-0.199	0.088
	(0.879)	(0.960)	(0.674)	(0.647)	(0.846)	(0.500)	(0.094)	(0.457)	(0.443)	(0.910)
Observations	112	112	112	112	112	112	112	112	112	112

Note: Poisson regressions. We use five dependent variables: *Adopt, Adopt New*, *Adopt Now, Adopt Delay, and Adopt Used. Adopt* counts the total number of technology adoption decisions. *Adopt New* counts only the number of technologies adopted that have never been used before. *Adopt Now* counts the total number of times an individual adopts a new technology in the round where it becomes available. *Adopt Delay* measures the number of times a new technology is adopted at least one round after it arrives. *Adopt Used* counts the number of times a previously used technology is adopted. Columns 1-5 use data on only common technologies between the two treatments (Baseline Technology through Technology J). Columns 6-10 use data on only the first five technologies. Robust standard errors. P-values in parentheses. Significance denoted by *** p<0.01, ** p<0.05, * p<0.1, respectively.

Table 3.7 presents the APEs for our five count variables using two different specifications: common technologies (Baseline through Technology J) and the first five technologies (Baseline through Technology D).[45] Coefficients for the count models are presented in appendix Table 3B.2. Beginning with the estimates for common technologies (columns 1-5 in Table 3.7), we find evidence that an increase in the rate of innovation decreases the number of technologies adopted for all adoption measures. Individuals in the high innovation treatment adopt 2.5 fewer technologies ($p<0.01$) overall through the course of the experiment. Decomposing these effects, in the high innovation treatment our results show the number of new technologies adopted declines by 1.3 technologies ($p<0.01$) – with individuals adopting 0.7 fewer technologies ($p<0.1$) in the period when they first are made available and 0.5 fewer technologies ($p<0.1$) with some level of delay. We also find higher rates of innovation reduces the number of used technologies adopted by 1.3 ($p<0.01$). Our findings provide strong support for (H2) where increases in the rate of innovation lead to a reduction in not only the total number of technologies adopted, but in the number of new technologies purchased in the experiment. This finding suggests that individuals are engaging in some delay by choosing to bypass certain innovations and wait to adopt future technologies that represent larger technological improvements once they arrive.

When we restrict our analysis to the first five technologies, we continue to find some (but relatively less) evidence that a higher rate of innovation motivates a lower number of technologies adopted. On average, individuals in the high innovation treatment adopt 1.1 fewer technologies ($p<0.05$). However, this effect is largely driven by a reduction in the number of used technologies adopted ($p<0.05$) and we find no statistically significant effect on new technology adoption.

[45] We exclude the comparison of all technologies because treatment status necessarily provides a much higher number of technologies (and therefore adoption decisions) over the course of the experiment. These restrictions ensure we are comparing the same technology sets across treatments.

Returning to our measure of risk aversion, we find the effects are constrained to the adoption of new innovations. For common technologies in Table 3.7 columns 1-5, we find that a one unit increase in the coefficient of relative risk aversion reduces the overall number of technologies adopted by an average of 1.3 ($p<0.1$), with most of this reduction coming from individuals adopting 0.8 fewer new technologies ($p<0.1$) in the round when a technology is first made available. This finding is consistent with earlier results, where more risk averse individuals are less likely to adopt a new innovation when it arrives.

3.6 Discussion and conclusions

Providing producers with both productivity enhancing agricultural technologies as well as the incentives to adopt technological improvements is an important part of a global strategy to maintain agricultural productivity growth rates. The rate at which new technologies become available to producers as well as the expected net return to a new innovation are important components of decisions over whether and when to adopt a costly new technology. Moreover, the uncertainty over future choices and returns generated by stochastic rates of innovation may generate incentives for individuals to delay adoption, especially when technological improvements are expected to arrive in the near future. Using a lab experiment, this article implements high and low innovation rate treatments to test for behavioral effects on the probability of adopting a new technology and the number of technologies adopted. We structure a decision environment where technologies generate stochastic returns, adoption requires payment of a fixed cost, and there are uncertain arrivals and returns to future innovations. We expected individuals in the high innovation treatment to be more likely to continue using a current technology in a given round and less likely to adopt a new technology as soon as it is made available **(H1)**; and to adopt fewer overall innovations when facing a common choice set **(H2)**. In addition to testing these hypotheses in the main experiment, we use a lottery choice experiment to measure subject risk aversion and test for its effects on individual adoption behavior.

Contrary to **H1**, our results suggest that, within a given round, exposure to a higher rate of innovation can increase the likelihood that an individual adopts a new technology in the first period it becomes available and reduce the probability that an individual continues to use a current technology. Additionally, the same individuals are less likely to adopt a new technology with delay or to switch back to a previously used technology in the high innovation treatment. This suggests that, on average, individuals in the high innovation treatment are more likely adopters of new innovations compared to the low innovation group.

One reason our findings might diverge from the predictions of the conceptual model could be biased expectations about the true rate of innovation in the low treatment group. One of the features of our innovation arrival pattern is a similar pattern of innovation during the first six rounds of the experiment. While subjects are provided with the actual expected innovation rate in the instructions (80% and 20% for the high and low innovation treatments, respectively), individuals could update their beliefs based on realized outcomes. For the low innovation group, four innovations arrive over the course of rounds two through six which corresponds to a rate of 80%, the same as the high innovation treatment, at the beginning of the experiment. This may have incentivized some individuals in the low innovation treatment to delay some adoption decisions for later arriving innovations when the observed rate of innovation arrival slowed down, contributing to the difference between treatments.[46]

[46] This pattern appears to hold for the low innovation treatment. In round 1, the average expected probability of a new innovation arrival in the next period is 38.7% and this average rises to 46.5% by round 5. Later in the experiment, once the observed rate of innovation slows down the average expected probability of arrival falls back

A second potential explanation for these findings is a heterogenous treatment effect given the parameterization of the experiment. Low switching costs combined with high returns can provide individuals with incentives to adopt improvements as soon as they arrive (Grenadier and Weiss 1997; Farzin et al. 1998). Individuals in the high innovation treatment may perceive the switching costs to be less of a burden when new technological improvements are arriving at a faster rate (especially as they earn more points in total during earlier rounds). In this experiment, the switching cost to adopt a new technology is relatively low, accounting for 62.5% and 26.9% of the lowest and highest expected per-period technology returns, respectively. More research with different fixed switching costs would be needed to further unpack this effect.

Although we generally reject **H1**, we do find strong evidence in support of **H2** that subjects in the high innovation rate environment adopt fewer new technologies than those in the low innovation rate environment when comparing across the same technology set (Baseline Technology through Technology J). When an individual adopts a new technology in the high innovation treatment, they then bypass the next few innovations. Even with a relatively short holding period, more new innovations are passed over by those in the high innovation treatment than in the low. This finding is consistent with patterns demonstrated via other simulation models (Grenadier and Weiss 1997). In the low innovation treatment, the strategy appears to be one where smaller technological improvements are adopted, albeit with some level of delay. It is also important to recall that although we constrain this analysis so that subjects in both treatment groups are faced with the same set of new technologies (Baseline Technology through Technology J), these technologies arrive over the course of 15 rounds for the high innovation treatment and 22 rounds for the low innovation treatment. Thus, we must also consider the time dimension, and the

---

to initial levels of 37.8%. Among the high innovation treatment, the average expected probability of a new innovation arrival is consistently close to the true value of 80%.

explicit pattern of innovation arrivals (or non-arrivals) as part of the treatment. Changing the pattern of arrival while keeping the overall innovation rate constant, could impact the results to the extent that individuals make adoption decisions based on innovation realizations (e.g. number of periods without an innovation).

Together, our two main findings describe a pattern of adoption behavior not altogether inconsistent with our conceptualization of delayed adoption. Individuals exposed to an increased rate of innovation appear to be more likely to adopt a new innovation as soon as it arrives. However, over time, these same individuals adopt fewer technologies over the course of the experiment. This suggests that in the high innovation rate environment, individuals are waiting for the arrival of new technologies, and when a sufficiently improved technology is made available, are more likely to immediately adopt it rather than delaying the decision any further. Future research should take into account the magnitude of technological improvements observed when adoption decisions occur to better understand how the rate of innovation might impact the optimal level of improvement individuals are looking for in a new technology.

Regarding the effect of subjects' level of relative risk aversion on their adoption behavior, we find that more risk averse individuals are less likely to adopt new innovations as soon as they arrive and are more likely to continue using current technologies. Increased levels of risk aversion also result in the adoption of fewer new technologies over time, consistent with other studies of timing (Carletto et al. 1999; Carletto et al. 2010). We find little evidence that risk aversion plays any role in the adoption of technologies with delay or switching back to old technologies. This suggests that risk averse individuals might be particularly sensitive to the uncertainty surrounding the arrival of *new* innovations and is largely consistent with descriptions of late adopters compared to early adopters (Rogers 1962).

This work has several implications for policymakers. First, it suggests that increasing the rate of technological innovation through higher public (and private) commitments to R&D spending as recommended by Pardey et al. (2015) is unlikely to have the unintended consequence of incentivizing significant delays in adoption. Instead, if faster rates of innovation make individuals more likely adopters of new technologies as soon as they become available, then maintaining a constant pipeline of new technologies is imperative for long-run productivity growth. Second, higher rates of innovation may result in fewer technologies being adopted in the short-run, as individuals bypass some incremental technological improvements in the favor of future gains. Policymakers wishing to incentivize the uptake of more, possibly smaller, innovations may well want to focus on reducing the costs of switching or releasing technologies with more significant improvements over those currently in use. Policies to lower adoption costs might include information dissemination via extension, public-private partnerships to reduce search costs and improve access in retail outlets, or even targeted technology subsidies for individuals with binding liquidity constraints.

This research also highlights several areas for future research. One is the need for the development of numerical simulation models that incorporate additional features of agricultural technology markets and innovation processes. In addition, incorporating features such as stochastic rates of innovation, depreciation, salvage value, and competitive technology pricing will help to develop further hypotheses about individual adoption behavior. Another area for further experimental extensions is to introduce agricultural decisionmakers into this controlled environment to see if adoption behavior diverges from that of student subjects. Additionally, exploring alternative parameterization of the experiment, especially in terms of the fixed adoption costs, may yield results more in line with real options models by raising the barrier to switching.

Finally, much more empirical work is needed to assess the impact of different innovation rates on various agricultural technologies. More detailed data collection on the types of innovations adopted (e.g. crop variety, equipment model numbers, software versions, etc.) rather than binary indicators for technology use will be required to evaluate the interplay of innovation patterns and individual behavior at a larger scale.

APPENDICES

APPENDIX 3A: Experimental instructions

Below we present the neutrally framed experimental instructions for our two innovation treatments. The only difference between the two sets of instructions is the stated probability of a new innovation arriving in a round.

High Innovation Treatment:

1. In the following experiment, you will be tasked with choosing technologies to maximize earnings from a generic production process over 26 rounds. Your earnings in each period will depend on the decisions that you make.

2. At the beginning of the experiment, you will be provided with an endowment of 100 points. These points can be used to purchase production technologies from a list of options that will become available during the experiment. In each round of the experiment, you get to choose which technology you want to use.

3. For each technology, you will see information regarding the average return generated by the technology as well as the range of possible returns. You have an equal chance of receiving any return in the range of possible returns during each period.

   For example, you may be presented with the following:

Technology B: Average return of 55 points with range between 20 and 90 points per period. Returns have a 70 point spread.

Choosing this technology means that you will have an equal chance of receiving any return between 20 and 90 points in any period. For example, the odds of this technology returning 25 points are the same as the odds of this technology returning 70 points. The spread is the range of possible returns. In this example, the range of possible returns you can receive is 70 points, the difference between the highest and lowest possible values (90-20=70).

4. At the end of each period, we will provide you with information regarding your point earnings based on the technology you chose to use. These point earnings will be recorded and added to your cumulative earnings.

5. At the beginning of each round, there is an **<u>80% chance</u>** that a new technology becomes available on the market and will be added to your list of technology options.

6. Should you choose to use a new technology that you have never used before, you will pay 50 points to purchase the new technology. Should you choose to switch back to a technology that you have used before, you will pay 25 points to switch back to that technology. There is no cost to continuing to use your current technology. There is no restriction on how many times you can change technologies, other than you must have enough points.

7. At the end of the experiment we will add up your point earnings and convert them to U.S. dollars at a rate of $1.00 per 500 points. We will pay you this amount in the form of a gift card emailed to the address you specify later. Your earnings are your own business and you do not have to discuss them with anyone.

Low Innovation Treatment:

1. In the following experiment, you will be tasked with choosing technologies to maximize earnings from a generic production process over 26 rounds. Your earnings in each period will depend on the decisions that you make.

2. At the beginning of the experiment, you will be provided with an endowment of 100 points. These points can be used to purchase production technologies from a list of options that will become available during the experiment. In each round of the experiment, you get to choose which technology you want to use.

3. For each technology, you will see information regarding the average return generated by the technology as well as the range of possible returns. You have an equal chance of receiving any return in the range of possible returns during each period.

   For example, you may be presented with the following:

Technology B: Average return of 55 points with range between 20 and 90 points per period. Returns have a 70 point spread.

Choosing this technology means that you will have an equal chance of receiving any return between 20 and 90 points in any period. For example, the odds of this technology returning 25 points are the same as the odds of this technology returning 70 points. The spread is the range of possible returns. In this example, the range of possible returns you can receive is 70 points, the difference between the highest and lowest possible values (90-20=70).

4. At the end of each period, we will provide you with information regarding your point earnings based on the technology you chose to use. These point earnings will be recorded and added to your cumulative earnings.

5. At the beginning of each round, there is an <u>**20% chance**</u> that a new technology becomes available on the market and will be added to your list of technology options.

6. Should you choose to use a new technology that you have never used before, you will pay 50 points to purchase the new technology. Should you choose to switch back to a technology that you have used before, you will pay 25 points to switch back to that technology. There is no cost to continuing to use your current technology. There is no restriction on how many times you can change technologies, other than you must have enough points.

7. At the end of the experiment we will add up your point earnings and convert them to U.S. dollars at a rate of $1.00 per 500 points. We will pay you this amount in the form of a gift card emailed to the address you specify later. Your earnings are your own business and you do not have to discuss them with anyone.

APPENDIX 3B: Robustness checks

Table 3B.1: Lottery choice experiment decisions and outcomes

Decision	Option A	Option B
	10% chance of 200 points	10% chance of 385 points
1	90% chance of 160 points	90% chance of 10 points
	20% chance of 200 points	20% chance of 385 points
2	80% chance of 160 points	80% chance of 10 points
	30% chance of 200 points	30% chance of 385 points
3	70% chance of 160 points	70% chance of 10 points
	40% chance of 200 points	40% chance of 385 points
4	60% chance of 160 points	60% chance of 10 points
	50% chance of 200 points	50% chance of 385 points
5	50% chance of 160 points	50% chance of 10 points
	60% chance of 200 points	60% chance of 385 points
6	40% chance of 160 points	40% chance of 10 points
	70% chance of 200 points	70% chance of 385 points
7	30% chance of 160 points	30% chance of 10 points
	80% chance of 200 points	80% chance of 385 points
8	20% chance of 160 points	20% chance of 10 points
	90% chance of 200 points	90% chance of 385 points
9	10% chance of 160 points	10% chance of 10 points
	100% chance of 200 points	100% chance of 385 points
10	0% chance of 160 points	0% chance of 10 points

Note: We enforced a single switching point in the above experiment which enforces consistency of preferences. At the end of the Stage 1 experiment, a lottery choice decision was randomly chosen, and individuals received the number of points from the execution of the lottery added to their final payment.

Table 3B.2: Coefficient estimates for the number of new technologies adopted (Poisson)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Sample	Common Technologies					First Five Technologies				
Dependent Variable	Adopt	Adopt New	Adopt Now	Adopt Delay	Adopt Used	Adopt	Adopt New	Adopt Now	Adopt Delay	Adopt Used
High innovation treatment	-0.413***	-0.257***	-0.142*	-0.475*	-1.128***	-0.350***	-0.106	-0.017	-0.502	-1.263***
	(0.001)	(0.002)	(0.084)	(0.060)	(0.009)	(0.010)	(0.198)	(0.808)	(0.133)	(0.006)
Relative risk aversion coefficient	-0.220*	-0.156*	-0.160**	-0.011	-0.420	-0.142	-0.030	-0.032	0.073	-0.455
	(0.092)	(0.066)	(0.048)	(0.965)	(0.367)	(0.311)	(0.666)	(0.619)	(0.820)	(0.358)
Age	0.051	0.016	-0.010	0.118	0.209	0.031	0.015	-0.020	0.213**	0.055
	(0.444)	(0.744)	(0.840)	(0.187)	(0.295)	(0.648)	(0.705)	(0.564)	(0.048)	(0.828)
Gender (1=Male)	-0.485***	-0.258***	-0.061	-0.899***	-2.065***	-0.434***	-0.145*	0.030	-1.291***	-1.973***
	(0.000)	(0.004)	(0.485)	(0.000)	(0.000)	(0.000)	(0.091)	(0.670)	(0.002)	(0.000)
Michigan resident (1=Yes)	-0.027	-0.020	-0.168	0.813**	0.225	0.114	-0.039	-0.192*	14.966***	1.145**
	(0.900)	(0.889)	(0.248)	(0.028)	(0.654)	(0.575)	(0.805)	(0.099)	(0.000)	(0.038)
Sophomore	0.429	0.053	0.075	-0.090	1.662*	0.194	-0.230	-0.229	0.450	1.434
	(0.248)	(0.856)	(0.796)	(0.846)	(0.058)	(0.559)	(0.338)	(0.288)	(0.496)	(0.166)
Junior	0.033	0.051	-0.026	0.321	0.038	-0.111	0.008	-0.048	0.370	-0.497
	(0.911)	(0.811)	(0.902)	(0.535)	(0.965)	(0.707)	(0.966)	(0.773)	(0.545)	(0.602)
Senior	-0.139	-0.104	-0.080	-0.101	-0.181	-0.297	-0.199	-0.138	-0.174	-0.423
	(0.587)	(0.513)	(0.624)	(0.829)	(0.810)	(0.243)	(0.138)	(0.320)	(0.777)	(0.591)
Graduate student	-0.037	0.008	0.071	-0.269	-0.132	-0.198	-0.232	-0.102	-0.570	0.093
	(0.879)	(0.960)	(0.670)	(0.635)	(0.844)	(0.496)	(0.104)	(0.457)	(0.391)	(0.909)
Constant	1.168	1.538	2.084*	-2.679	-3.685	0.926	0.846	1.811**	-19.760***	-1.422
	(0.446)	(0.178)	(0.066)	(0.180)	(0.415)	(0.541)	(0.371)	(0.022)	(0.000)	(0.804)
Observations	112	112	112	112	112	112	112	112	112	112

Note: Poisson regressions. We use five dependent variables: *Adopt, Adopt New*, *Adopt Now, Adopt Delay, and Adopt Used. Adopt* counts the total number of technology adoption decisions. *Adopt New* counts only the number of technologies adopted that have never been used before. *Adopt Now* counts the total number of times an individual adopts a new technology in the round where it becomes available. *Adopt Delay* measures the number of times a new technology is adopted at least one round after it arrives. *Adopt Used* counts the number of times a previously used technology is adopted. Columns 1-5 use data on only common technologies between the two treatments (Baseline Technology through Technology J). Columns 6-10 use data on only the first five technologies. Robust standard errors. P-values in parentheses. Significance denoted by *** p<0.01, ** p<0.05, * p<0.1, respectively.

APPENDIX 3C: Duration of adoption

We also analyze the duration of average adoption periods directly, by considering the number of rounds individuals spend using a technology or engaging in delay. In Table 3C.1, Panel A we compare the difference in means for the high and low innovation treatments to evaluate if there is a significant effect on the amount of time between technology changes.[47] We again cover several different technology measures including unconditional adoption (*Adopt*), adopting any new technology with or without delay (*Adopt New*), adoption without delay (*Adopt Now*), adoption with delay (*Adopt Delay*), and adopting a used technology (*Adopt Used*). Using data on all technologies in the experiment, we find no evidence of significant differences in the average time between different adoption events. However, when we focus only on technologies common to both studies we see that individuals in the high innovation treatment switch technologies 1.75 rounds sooner than those in the low innovation treatment ($p<0.01$). This effect persists when we only consider the adoption of new technologies in total as well as for new technologies adopted with and without delay. These findings are largely consistent when we restrict the analysis to the first five technologies.

While the results provide some evidence that individuals in the high innovation environment spend less time between technology adoption events, this could still be a function of the spacing of innovation arrivals. In Table 3C.1, Panel B we test for differences in length of technology delay between treatment groups. Specifically, we measure the average number of rounds an individual delays the adoption of a new innovation, conditional on eventually adopting it. For example, if an individual adopts a technology as soon as it is made available we would

[47] Our between-subjects design does not permit the estimation of a duration model (survival analysis) for the data we collected here for two related reasons. First, we cannot control for differences in the underlying utility function across individuals who are only assigned to a single treatment group. Second, the hazard of switching technologies is fundamentally different across our two treatments due to the stochastic arrival and returns to each technology.

measure the delay as zero. If instead an individual adopts a technology three rounds after it first enters the choice set, delay would be coded as three rounds. Comparing all technologies, we find that subjects in the high innovation group delay adoption decisions by an average of 0.47 rounds, which is significantly less than the 1.43 round average delay exhibited in the low innovation group ($p<0.01$). We find a similar effect when comparing the average delay among the common technologies where the average delay in the high innovation treatment is 0.99 rounds shorter than in the low innovation treatment ($p<0.01$). We find no evidence of a difference among the first five technologies.

Table 3C.1: Average time using a technology and average delay

Panel A: Average time between technology adoption decisions

	Average Time Between Adoption				
	High	Low	Difference	Mann-Whitney	Two-tailed t-test
All technologies					
Adopt	5.564	5.568	-0.004	0.710	0.994
Adopt New	5.796	5.596	0.200	0.512	0.756
Adopt Now	7.155	7.317	-0.162	0.746	0.828
Adopt Delay	12.063	9.486	2.577	0.189	0.078
Adopt Used	6.235	6.604	-0.369	0.475	0.833
Common technologies					
Adopt	3.813	5.564	-1.751***	0.001	0.000
Adopt New	4.058	5.956	-1.898***	0.000	0.000
Adopt Now	4.420	7.317	-2.897***	0.000	0.000
Adopt Delay	6.620	9.486	-2.866**	0.015	0.014
Adopt Used	6.320	6.604	-0.284	0.447	0.871
First five technologies					
Adopt	2.392	2.712	-0.320*	0.099	0.167
Adopt New	2.534	2.787	-0.253**	0.036	0.429
Adopt Now	2.112	2.671	-0.559***	0.005	0.006
Adopt Delay	4.081	3.944	0.137	0.258	0.885
Adopt Used	5.910	6.968	-1.058	0.999	0.617

Note: We measure five types of adoption durations: *Adopt, Adopt New, Adopt Now, Adopt Delay,* and *Adopt Used*. *Adopt* is the average number of rounds an individual uses a technology before switching. *Adopt New* is the average number of rounds using a new technology before switching. *Adopt Now* is the average number of rounds between adopting a new technology as soon as it arrives. *Adopt Delay* is the average number of rounds between adopting a new technology with at least a one round delay. *Adopt Used* is the average number of rounds between adopting a used technology. We report p-values for the nonparametric Mann-Whitney test and the two-tailed t-test respectively. Significance denoted by *** p<0.01, ** p<0.05, * p<0.1, respectively, using the Mann-Whitney results. We drop all durations that are right-censored (have not ended by round 26) to standardize comparisons across treatments.

Panel B: Average delay for technology adoption

	Average Delay				
Sample	High	Low	Difference	Mann-Whitney	Two-tailed t-test
All technologies	0.472	1.431	-0.959***	0.009	0.000
Common technologies	0.444	1.431	-0.987***	0.001	0.000
First five technologies	0.347	0.518	-0.171	0.421	0.379

Note: Delay is measured as the number of rounds individuals delay adoption of chosen technologies. A technology adopted in the period it is made available is recorded as a delay of zero rounds. We report p-values for the nonparametric Mann-Whitney test and the two-tailed t-test respectively. Significance denoted by *** p<0.01, ** p<0.05, * p<0.1, respectively, using the Mann-Whitney results. We drop all durations that are right-censored (have not ended by round 26) to standardize comparisons across treatments.

REFERENCES

REFERENCES

Abdulai, A., and W.E. Huffman. 2005. "The Diffusion of New Agricultural Technologies: The Case of Crossbred-Cow Technology in Tanzania." *American Journal of Agricultural Economics* 87(3):645-.

Alston, J.M., M.A. Andersen, J.S. James, and P.G. Pardey. 2010. *Persistence pays: U.S. agricultural productivity growth and the benefits from public R&D spending*. New York: Springer.

Alston, J.M., M.A. Andersen, J.S. James, and P.G. Pardey. 2011. "The Economic Returns to U.S. Public Agricultural Research." *American Journal of Agricultural Economics* 93(5):1257–1277.

Alston, J.M., C. Chan-Kang, M.C. Marra, P.G. Pardey, and T. Wyatt. 2000. *A Meta-Analysis of Rates of Return to Agricultural R&D:Ex Pede Herculem?* International Food Policy Research Institute. Available at: http://ebrary.ifpri.org/cdm/ref/collection/p15738coll2/id/125334 [Accessed May 24, 2018].

Baerenklau, K.A., and K.C. Knapp. 2007. "Dynamics of Agricultural Technology Adoption: Age Structure, Reversibility, and Uncertainty." *American Journal of Agricultural Economics* 89(1):190-.

Balcer, Y., and S.A. Lippman. 1984. "Technological expectations and adoption of improved technology." *Journal of Economic Theory* 34(2):292–318.

Ball, E., D. Schimmelpfennig, and S.L. Wang. 2013. "Is U.S. Agricultural Productivity Growth Slowing?" *Applied Economic Perspectives and Policy* 35(3):435–450.

Biggs, S.D. 1990. "A multiple source of innovation model of agricultural research and technology promotion." *World Development* 18(11):1481–1499.

Bruhn, M., and D. McKenzie. 2009. "In Pursuit of Balance: Randomization in Practice in Development Field Experiments." *American Economic Journal: Applied Economics* 1(4):200–232.

Burton, M., D. Rigby, and T. Young. 2003. "Modelling the adoption of organic horticultural technology in the UK using Duration Analysis." *The Australian Journal of Agricultural and Resource Economics* 47(1):29–54.

Cai, Y., A.A. Golub, and T.W. Hertel. 2017. "Agricultural research spending must increase in light of future uncertainties." *Food Policy* 70:71–83.

Carey, J.M., and D. Zilberman. 2002. “A Model of Investment under Uncertainty: Modern Irrigation Technology and Emerging Markets in Water.” *American Journal of Agricultural Economics* 84(1):171–183.

Carletto, C., A. de Janvry, and E. Sadoulet. 1999. “Sustainability in the Diffusion of Innovations: Smallholder Nontraditional Agro-Exports in Guatemala.” *Economic Development and Cultural Change* 47(2):345–369.

Carletto, C., A. Kirk, P.C. Winters, and B. Davis. 2010. “Globalization and Smallholders: The Adoption, Diffusion, and Welfare Impact of Non-Traditional Export Crops in Guatemala.” *World Development* 38(6):814–827.

Caswell, M., and D. Zilberman. 1985. “The Choices of Irrigation Technologies in California.” *American Journal of Agricultural Economics* 67(2):224.

Chen, D.L., M. Schonger, and C. Wickens. 2016. “oTree—An open-source platform for laboratory, online, and field experiments.” *Journal of Behavioral and Experimental Finance* 9:88–97.

Chen, M., and R.J. Myers. 2017. “Understanding Transient Technology Use among Smallholder Farmers in Africa.” *AgEcon Search*.

Conley, T., and C. Udry. 2010. “Learning about a New Technology: Pineapple in Ghana.” *The American Economic Review* 100(1):35–69.

Dixit, A.K., and R.S. Pindyck. 1994. *Investment under uncertainty*. Princeton, N.J: Princeton University Press.

Dong, D., and A. Saha. 1998. “He came, he saw, (and) he waited: an empirical analysis of inertia in technology adoption.” *Applied Economics* 30(7):893–905.

Doraszelski, U. 2004. “Innovations, improvements, and the optimal adoption of new technologies.” *Journal of Economic Dynamics and Control* 28(7):1461–1480.

Doraszelski, U. 2001. “The net present value method versus the option value of waiting: A note on Farzin, Huisman and Kort (1998).” *Journal of Economic Dynamics and Control* 25:7.

Duflo, E., M. Kremer, and J. Robinson. 2011. “Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya.” *American Economic Review* 101(6):2350–2390.

Emerick, K., A. de Janvry, E. Sadoulet, and M.H. Dar. 2016. “Technological Innovations, Downside Risk, and the Modernization of Agriculture.” *American Economic Review* 106(6):1537–1561.

Farzin, Y.H., K.J.M. Huisman, and P.M. Kort. 1998. “Optimal timing of technology adoption.” *Journal of Economic Dynamics and Control* 22(5):779–799.

Feder, G., R.E. Just, and D. Zilberman. 1985. "Adoption of Agricultural Innovations in Developing Countries: A Survey." *Economic Development and Cultural Change* 33(2):255–298.

Feder, G., and D.L. Umali. 1993. "The adoption of agricultural innovations." *Technological Forecasting and Social Change* 43(3–4):215–239.

Foster, A., and M. Rosenzweig. 1995. "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture." *Journal of Political Economy* 103(6):1176–1209.

Foster, A.D., and M.R. Rosenzweig. 2010. "Microeconomics of Technology Adoption." *Annual Review of Economics*:395–424.

Fudenberg, D., and J. Tirole. 1985. "Preemption and Rent Equalization in the Adoption of New Technology." *The Review of Economic Studies* 52(3):383.

Giné, X., and D. Yang. 2009. "Insurance, credit, and technology adoption: Field experimental evidence from Malawi." *Journal of Development Economics* 89(1):1–11.

Greiner, B. 2015. "Subject pool recruitment procedures: organizing experiments with ORSEE." *Journal of the Economic Science Association* 1(1):114–125.

Grenadier, S.R., and A.M. Weiss. 1997. "Investment in technological innovations: An option pricing approach." *Journal of Financial Economics* 44(3):397–416.

Harou, A.P., T.F. Walker, and C.B. Barrett. 2017. "Is late really better than never? The farmer welfare effects of pineapple adoption in Ghana." *Agricultural Economics* 48(2):153–164.

Holt, C.A., and S.K. Laury. 2002. "Risk Aversion and Incentive Effects." *The American Economic Review* 92(5):1644–1655.

Hoppe, H.C. 2000. "Second-mover advantages in the strategic adoption of new technology under uncertainty." *International Journal of Industrial Organization* 18(2):315–338.

Hoppe, H.C. 2002. "The Timing of New Technology Adoption: Theoretical Models and Empirical Evidence." *The Manchester School* 70(1):56–76.

Huh, Y.E., and S.-H. Kim. 2008. "Do early adopters upgrade early? Role of post-adoption behavior in the purchase of next-generation products." *Journal of Business Research* 61(1):40–46.

Huisman, K.J.M., and P.M. Kort. 2004. "Strategic technology adoption taking into account future technological improvements: A real options approach." *European Journal of Operational Research* 159(3):705–728.

Hurley, T.M., X. Rao, and P.G. Pardey. 2014. "Re-examining the Reported Rates of Return to Food and Agricultural Research and Development." *American Journal of Agricultural Economics* 96(5):1492–1504.

Isik, M., and M. Khanna. 2003. "Stochastic Technology, Risk Preferences, and Adoption of Site-Specific Technologies." *American Journal of Agricultural Economics* 85(2):305–317.

de Janvry, A., and J.J. Dethier. 1985. "Technological Innovation in Agriculture: The Political Economy of its Rate and Bias." *Consultative Group on International Agricultural Research (CGIAR)*. Available at: http://link.springer.com/10.1007/BF02772888 [Accessed June 18, 2018].

Jensen, R. 1982. "Adoption and diffusion of an innovation under uncertain profitability." *Journal of Economic Theory* 27(1):182–199.

Jensen, R. 1988. "Information Cost and Innovation Adoption Policies." *Management Science* 34(2):230–239.

Karlan, D., R. Osei, I. Osei-Akoto, and C. Udry. 2014. "Agricultural Decisions after Relaxing Credit and Risk Constraints." *The Quarterly Journal of Economics* 129(2):597–652.

Kilelu, C.W., L. Klerkx, and C. Leeuwis. 2013. "Unravelling the role of innovation platforms in supporting co-evolution of innovation: Contributions and tensions in a smallholder dairy development programme." *Agricultural Systems* 118:65–77.

Kim, S.-H., and V. Srinivasan. 2009. "A Conjoint-Hazard Model of the Timing of Buyers' Upgrading to Improved Versions of High-Technology Products." *Journal of Product Innovation Management*:278–290.

McCardle, K.F. 1985. "Information Acquisition and the Adoption of New Technology." *Management Science* 31(11):1372–1389.

McDonald, R., and D. Siegel. 1986. "The Value of Waiting to Invest." *The Quarterly Journal of Economics* 101(4):707–728.

Moser, C.M., and C.B. Barrett. 2003. "The disappointing adoption dynamics of a yield-increasing, low external-input technology: the case of SRI in Madagascar." *Agricultural Systems* 76(3):1085–1100.

Mukherji, N., B. Rajagopalan, and M. Tanniru. 2006. "A decision support model for optimal timing of investments in information technology upgrades." *Decision Support Systems* 42(3):1684–1696.

Munshi, K. 2004. "Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution." *Journal of Development Economics* 73(1):185–213.

Neef, A., and D. Neubert. 2011. "Stakeholder participation in agricultural research projects: a conceptual framework for reflection and decision-making." *Agriculture and Human Values* 28(2):179–194.

Pardey, P.G., C. Chan-Kang, J.M. Beddow, and S.P. Dehmer. 2015. "Long-run and Global R&D Funding Trajectories: The U.S. Farm Bill in a Changing Context." *American Journal of Agricultural Economics* 97(5):1312–1323.

Purvis, A., W.G. Boggess, C.B. Moss, and J. Holt. 1995. "Technology Adoption Decisions under Irreversibility and Uncertainty: An Ex Ante Approach." *American Journal of Agricultural Economics* 77(3):541–551.

Reinganum, J.F. 1981. "On the Diffusion of New Technology: A Game Theoretic Approach." *The Review of Economic Studies* 48(3):395.

Rogers, E.M. 1962. *Diffusion of innovations* 1st ed. New York: Free Press.

Smirnov, V., and A. Wait. 2015. "Innovation in a generalized timing game." *International Journal of Industrial Organization* 42:23–33.

Spielman, D. 2005. "Innovation Systems Perspectives on Developing-Country Agriculture: A Critical Review." ISNAR Discussion Paper No. 2, International Food Policy Research Institute.

Suri, T. 2011. "Selection and Comparative Advantage in Technology Adoption." *Econometrica* 79(1):159–209.

Tjernström, E. 2015. "Signals, Similarity and Seeds: Social Learning in the Presence of Imperfect Information and Heterogeneity∗." *Unpublished Manuscript*:42.

Wang, S.L., P. Heisey, D. Schimmelpfennig, and E. Ball. 2015. "Agricultural Productivity Growth in the United States: Measurement, Trends, and Drivers." *Economic Research Service Economic Research Report 189*:78.

Weiss, A.M. 1994. "The Effects of Expectations on Technology Adoption: Some Empirical Evidence." *The Journal of Industrial Economics* 42(4):341.

Yoon, Y.-R. 2009. "Endogenous timing of actions under conflict between two types of second mover advantage." *International Journal of Industrial Organization* 27(6):728–738.