THE FUTURE AGENDA OF EXPERIMENTAL ECONOMICS: INNOVATIONS, REAL-WORLD APPLICATIONS AND INTERDISCIPLINARY DIRECTIONS

Rakesh Kumar Sudan

Associate Professor, University College, Kurukshetra, India

E-mail: duddyx@gmail.com

ABSTRACT

A desirable agenda for experimental economics involves exploring new areas within behavioral economics, developing innovative experimental designs and applications in real-world settings, and fostering interdisciplinary collaborations to address complex economic questions. Key future directions include providing readers with potential avenues for research, testing the validity of economic theories, and using experimental insights to design better markets and institutions. The paper delves into the contributions of well-established researchers in economics and psychology about the viability of experimental data in labs and field settings and the inferences made out.

Keywords: *Credibility Inflexion Point; Causal Effect; Randomized Assignment; Juvenile Delinquency; Science of Scaling; Psychology's Renaissance.*

INTRODUCTION

The past few years have seen a veritable explosion of randomized experiments in development economics. At the fall 2008 NEUDC (Northeast Universities Development Consortium) conference, a large conference in development economics attended mainly by young researchers and PhD students, 24 papers reported on randomized field experiments, out of the 112 papers that used microeconomics data (laboratory experiments excluded). This is up from 4 in 2004. At the fall 2008 BREAD (Bureau for Research and Development in Economic Analysis) conference, the premier conference on development economics, 4 of the 8 papers invited were randomized field experiments. Three out of the 6 papers using microeconomic data from developing countries published in 2008 or forthcoming in the Quarterly Journal of Economics involve randomized assignment. The enthusiasm is not limited to academia. At the World Bank, there are 67 ongoing randomized evaluations (out of 89 ongoing pro-gram evaluations) in the African region alone. Perhaps inevitably, this progress has also generated a rising tide of criticism. Almost all of the criticism is well meant, recognizing the benefits of such experiments while suggesting that we not forget that there are a lot of important questions that randomized experiments cannot answer.

Much of this criticism has been useful even when we do not entirely agree with it, both in helping us define the strengths and limitations of randomized experiments, and in clarifying where the field needs to go next. However, we argue that much of this criticism misses (or at least insufficiently emphasizes) the main reasons why there has been so much excitement surrounding experimental research in development economics. We then return to the various criticisms, in part to clarify and qualify them and to argue that, because of an imperfect recognition of what is exciting about the experimental agenda, there is a tendency to set up false oppositions between experimental work and other forms of research.

THE PROMISE OF EXPERIMENTS

Experimental research in economics, like earlier research in health, and education, started from a concern about the reliable identification of program effects in the face of complex and multiple channels of causality. In general, participants to a program differ from nonparticipants in many ways, and we have no information on how they would have fared had they not been participating. This makes it difficult to separate the "causal effect" of the program (i.e., for a given participant, the difference between the outcome he experienced under the program and the outcome he would have experienced if he had not been participating) from other factors. A central problem is selection, the fact that participants may be systematically different from nonparticipants. Although the treatment effect for each person cannot be identified, experiments make it possible to vary one factor at a time and therefore provide internally valid estimates of the average treatment effect for a population of interest.

EXPERIMENTAL ECONOMICS IN PHILOSOPHICAL PERSPECTIVE

The experimental science community of social sciences finds itself at crossroads of 'credibility inflexion point' in matter of experimental research. Experiments in constrained settings take away, more or less, the realism of the observations. Quality of existing data, data gaps is also an issue. Thus experimental necessity of social sciences becomes a buzzword in the present advanced technology of data analysis. However certain issues like replicability of the experiments hang around in the experimental domain related to economics.

However, it's not just replicability issue that lies at the core. While it is of first-order importance to embrace research practices that increase reproducibility, experimental economists also must tackle the generalizability and applicability of the evidence they produce. For example, permanent income is an important explanatory variable in Friedman's consumption-income relationship.

Many ways can guide us to an estimate of this entity. However, one which gives the best approximation needs to be explored. Same can be said about the factors determining the shape of Phillips curve After all, we are not only interested in ensuring that the same experiment yields the same outcome when it is repeated. Ideally, we would like to be able to generalize our findings to different contexts and produce insights that contribute to economic theory and policy discussions.

These considerations have given rise a constructive debate that is pinned on the credibility, generalizability, and relevance of findings in experimental economics. Several recommendations are put up here that summarize research practices that we as researchers should all do more of. This is intended as an aid memoire or a ready reckoner for researchers in the design phase of their experiment.

While the guidelines relate to experimental economics in original, it is believed many of the suggestions are relevant in other fields of experimental science, and for scholars pursuing estimates using observational data. For example, how effective are exit polls in deducing about Voters' behavior how much child abuse stays as a dominant variable in explaining juvenile delinquency at a later stage.

It is desirable to cover four threats to generalizability: the representativeness of population, non-random selection into the experiment, treatment non-compliance, and characteristics of the experiment that may affect behavior. It is important that we do not merely acknowledge these issues ex post as potential limitations of our study. They must guide decisions in the design phase that affect the type of data we generate.

In principle, as well as in practice, there is good point in advocating for conducting more natural field experiments (NFEs). They are covert, and so can mitigate biases arising from self-selection into the experiment and experimenter demand effects, and typically involve the population of interest. NFEs offer a unique combination of control and realism. Also allied with NFEs is the ethical aspect. For example, households' economic behavior under conditions of sensitive issues like same-sex marriages, live-in relationships etc. pose challenges of such kind.

It can also be argued that lab and field experiments, as well as naturally occurring data, are complements in the production of scientific knowledge. For instance, we can begin by documenting an effect among students in the lab, then test its generalizability by repeating the experiment with different tasks and populations. Alternatively, we could first evaluate a program in a field experiment, then use additional lab experiments to test mechanisms that may explain what we observed in the field (Stoop 2012).

The second challenge is the informativeness of our findings – in other words, how to design experiments that optimize learning for the students and understanding of results at commonplace.

This requires a critical look at the practice of basing our inference solely on p-values, an approach that leads to a high false-positive rate (Ioannidis 2005) and ignores the economic significance of findings (Ziliak and McCloskey 2004). This practice is especially dangerous in combination with specification searching (Simmons 2011) and multiple hypothesis testing.

lt would be worthwhile to advocate for more replication studies to increase the credibility of research findings. We must apply our minds to devise ways to incentivize such studies. The incentives maybe in terms of fellowships, best research award or students' aid programs.

Due attention ought to be given to the importance of statistical power in determining the informative value of experiments, and present ways to increase power for a given experimental budget, such as using within-subject designs (where the same participant is exposed to different treatments, with their order randomized) when appropriate, and collecting baseline characteristics to perform blocked randomization (partitioning our sample to subgroups along relevant variables, and randomizing within these groups).

Finally, it would be appropriate to highlight issues related to the policy relevance of experimental economics results. Even perfectly credible and reproducible findings may not inform policy discussions if they exclusively focus on short-term impacts, leave mechanisms uncovered, and fail to consider scalability. Observing the longer-term outcomes ensures that promising results don't pass away quickly but is also important because it may take time for important general equilibrium effects to emerge due to the existence of inside and outside policy lags. For example, maybe an intervention led to short-term improvements for the participants, but in the long run transformed the market in ways that were harmful for everyone.

Experiments that document a resultant effect but do not explicitly study the underlying mechanisms leave a lot of potential gains on the table. Theory can help us address the generalizability threats discussed above by explicitly modelling the participation and compliance decision. Specifying a model, then using experimental variation to identify its deep structural parameters, allows us to extrapolate our results to different contexts and interventions. We can also design experiments to test the predictions of a theory, or to run a tiebreaker between competing models. If we can understand why we observe a phenomenon, we are in a better position to advise policy.

Theory also becomes important when we consider the 'science of scaling': a systematic treatment of the issues that arise when a small-scale, short-run program is rolled out on a much larger scale.

These considerations include general equilibrium effects (including the reaction of politicians to programs), and potential biases stemming from sample selection (the original pilot trial often includes a 'convenience' rather than a representative sample), site selection (researchers might choose to run their experiments in places where they are 'easier' to implement), and piloting (including the fact that as a program grows larger, it needs to recruit additional workers, who may be less skilled or motivated than the ones already hired. Researchers need to 'backward induct', to address scaling-related challenges in the design phase. If they do not, they may not create programs that work well at scale.

Grounds for optimism exist already as the researchers are becoming increasingly aware of these challenges and are taking important steps to improve the quality of their research. The profession as a whole is raising upto new standards of evidence (an example being the success of the preregistration movement).More academic work is needed for contributing to this positive change, to prompt other experimental researchers to join the debate, to share the challenges they have identified and their proposed suggestions to overcoming them, so that step-by-step we can work together to improve the quality of scientific research.

CONCLUSION

Many paradoxical situations arise in the realm of experimental economics, for that matter in other social sciences as well. Some are apparent, some are real ones. One solution is- Experimental domain of theory testing may be set aside from that of policy determining experiments.

REFERENCES

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine, 2*(8), 696–701. https://doi.org/10.1371/journal.pmed.0020124

Maniadis, Z., Tufano, F., & List, J. A. (2014). One swallow doesn't make a summer: New evidence on anchoring effects. *The American Economic Review, 104*(1), 277–290. https://doi.org/10.1257/aer.104.1.277

Stoop, J., Noussair, C. N., & Van Soest, D. (2012). From the lab to the field: Cooperation among fishermen. *Journal of Political Economy, 120*(6), 1027–1056. https://doi.org/10.1086/667809