Structured Ethics Appendices in Social Science Papers?
A Comment on Karlan and Udry's Proposal

By Douglas MacKay, PhD
UNC, Associate Professor

Dean Karlan and Christopher Udry have recently proposed that social scientists carrying out human subjects research append a "structured ethics appendix" to their papers. This appendix would consist of a set of questions which would "provide researchers with a concise and consolidated platform to spark thoughtful consideration of major ethical issues and to provide relevant information to readers (Karlan and Udry 2020)." Karlan and Udry rightly note that many studies raise ethical concerns that fall outside of the purview of institutional review boards (IRBs), and that these concerns ought to be considered and addressed in the planning of research, and publicly discussed to prevent readers from forming incorrect views. They provide a thoughtful example of a structured ethics appendix and illustrate its usefulness by applying it to four case studies.

Karlan and Udry pitch their appendix as a proposal, inviting feedback from the community. I take up their invitation in this post, offering a few comments on the questions included in the appendix with the aim of spurring conversation regarding what I take to be a timely and worthwhile endeavor.

Karlan and Udry's Proposed "Structured Ethics Appendix"

Karlan and Udry's proposed appendix includes 14 questions for researchers to answer. On the whole, these questions are well-considered and prompt researchers to consider the aspects of their studies that are most likely to raise ethical concerns left unaddressed by IRB review. Still, some of these questions could be revised to better capture the ethical issues at stake, and greater attention should be paid to the issue of informed consent.

Equipoise

Karlan and Udry's structured appendix asks researchers first to discuss whether their study satisfies the requirement of equipoise. They write:

> Clinical equipoise means there is genuine and meaningful uncertainty or disagreement amongst stakeholders on the outcome of the research (e.g., cost-effectiveness of an intervention relative to alternatives). Note this could be cut and paste from the main paper, the section in which one describes the existing literature on this topic and where this paper adds value to that literature (Karlan and Udry 2020).

Equipoise is indeed an important – though contested – norm of clinical research and I've argued elsewhere that it is relevant for policy research (see MacKay 2018 and MacKay 2020a). However, I wonder if it is best suited to capture the ethical concern Karlan and Udry have in mind here. As Karlan and Udry understand it, the defining feature of equipoise is that it denotes a state of uncertainty among stakeholders regarding the outcomes of the research. The research therefore has the potential to add "value" to the relevant literature (Karlan and Udry 2020). If Karlan and Udry's chief concern here is to ensure that studies have the potential to add value to ongoing debates, then it might be more suitable for the appendix to ask researchers to discuss the 'social value' of their research. Social value is widely recognized as an ethical requirement for clinical research and is satisfied when research promises to contribute to generalizable knowledge and/or evaluate diagnostic or therapeutic interventions that could lead to improvements in health or wellbeing (Emanuel, Wendler, and Grady 2000). By extension, policy research would count as socially valuable if, for example, it evaluated an untested intervention which promised to improve people's food security or literacy scores, or if it tested a prediction of a promising theory.

Equipoise, as Karlan and Udry define it above, is no doubt necessary for a study to be socially valuable, but this does not mean it is sufficient – a problem if their principal concern is to ensure that research contributes to a literature. More importantly, Karlan and Udry's understanding of equipoise overlooks an additional role it is often understood to play, namely, reconciling the obligation to promote participants' health interests with the practice of clinical research (Freedman 1987; Miller and Weijer 2003; and London Forthcoming). Clinical equipoise prohibits participants from being randomly assigned to an experimental intervention that is expected to be inferior to the standard of care treatment, thus ensuring participants' health interests are not sacrificed for the purposes of research. Similarly, in my work on the ethics of randomization, I've suggested that compliance with *policy equipoise* is one way to ensure that participants are treated fairly in policy research (MacKay 2020a; see also London Forthcoming). On this understanding, equipoise is satisfied when there is reasonable disagreement among the expert social scientific community regarding whether the intervention arm better promotes participants' interests compared to the 'standard policy.'

So, while Karlan and Udry are right that equipoise is relevant to the ethics of social science research, it is worth considering whether the social value requirement better captures their concern to ensure studies make a contribution to knowledge. In addition, because their understanding of equipoise is limited in the way I suggest above, it is also worth considering whether there is a role for a more robust conception of this concept in the appendix. I suggest that there is such a role below.

Informed Consent

Informed consent is a central protection for research participants. Karlan and Udry discuss it under question 2 of their appendix which concerns the role of researchers with respect to the implementation of the program. They write:

> Are researchers "active" researchers, i.e. did the researchers have direct decision-making power over whether and how to implement the program? If YES, what was the disclosure to participants and informed consent process for participation in the program? Providing IRB

> approval details may be sufficient, as in #2 above, but further clarification of any important issues should be discussed here. If NO, i.e., implementation was separate, explain the separation (Karlan and Udry 2020).

I agree with Karlan and Udry that the role of researchers is important and should be disclosed, but the issue of informed consent is deserving of greater attention in the appendix, perhaps as its own question. First, I worry that this way of posing the question assumes that if researchers are involved in deciding whether and how to implement a program, informed consent from participants is ethically necessary. While some commentators adopt this position (see Hoffman 2020), this claim is contentious, and many studies proceed without securing informed consent. For example, in the context of cluster randomized controlled trials, some argue that the requirement of informed consent may be waived if (1) the research would not be feasible without such a waiver; and (2) the study interventions and data collection procedures pose no more than minimal risk (Weijer et al. 2012). Averi Chakrabarti and I have also argued that since the primary function of informed consent is to respect people's autonomy rights, informed consent is not necessary in government policy experiments when the research does not infringe these rights. More specifically, we suggest that informed consent is not necessary when:

1. The government institution conducting or authorizing the research possesses a right to rule over the spheres of policy targeted by the research.
2. Data collection does not involve the infringement of participants' autonomy rights; and
3. Obtaining consent is impracticable (MacKay and Chakrabarti 2019).

I also think, second, that researchers should be required to discuss the informed process and provide a justification for any waiver of consent even if they did not have direct decision-making power over whether and how to implement the program. Researchers should be accountable for their decisions to partner with agencies in cases where those agencies unjustifiably waive the requirement to secure participants' informed consent. That is, if a government agency ignores its duty to secure participants' consent when carrying out a policy experiment, researchers analyzing the data don't have "clean hands" simply because they were not involved in the decisions regarding whether and how to carry out the experiment. Similarly, Karlan and Udry (2020) note in their discussion of their question regarding potential harms to participants that researchers should discuss such harms even if they had no role in designing or implementing the intervention.

To address these two problems, I would suggest that any structured appendix include something like the following question:

> Which information was disclosed to participants? If participants consented to participation in the study, how did the informed consent process proceed? If participants were not asked to consent to participation, what was the justification for waiving this requirement?

Scarcity

Karlan and Udry's (2020) question 7 addresses the ethics of randomization: "Did the inclusion of random assignment to treatment and/or control arms cause a change in the expected

aggregate value of programs or products delivered?" The ethical concern here is that people should not be denied access to a program for the purposes of conducting an RCT. Karlan and Udry's implicit proposal is that randomization is permissible when the program is scarce and the use of random assignment does not lead to less of the program in the aggregate.

I have two concerns with this question and the accompanying discussion. First, if there is genuine equipoise, such that the outcomes of the RCT are uncertain, what is the basis for the concern that the expected aggregate value of the programs is less in the context of an RCT? That is, if we don't know whether the intervention is superior to the control, in what sense is it worse to assign people to the control arm?

Second, while I agree that scarcity is relevant to the ethics of randomization when there is evidence to think that the intervention is superior to the control (MacKay 2020a), this doesn't mean that random assignment is permissible so long as the expected aggregate value is unchanged. Instead, for random assignment to be permissible, it must also be the case that a lottery is a fair way to allocate access to the program. The reason for this is that lotteries offer a just way to allocate a scarce good only when potential recipients have equally strong claims to the good; it is not just to use a fair lottery when some have a stronger claim to the good in question (MacKay 2020a). With respect to program delivery, there may be cases where some clients are in greater need of the program and so have a stronger claim to access it. For example, in the case of a cluster RCT, some communities might be worse off than others and so have a stronger claim to the program in question than others. Similarly, even in cases of randomization at the individual level, program officers might rightly object to allocating access to a scarce good – e.g. a housing voucher – by means of random assignment, since their clients differ in terms of the urgency of their need (MacKay 2020a).

The appropriate principle of allocation may not therefore be equal opportunity – i.e. a fair lottery – but rather priority to the worse off. A focus on aggregate value ignores the fact that *who* is served by a program, not merely the number of people served, may matter morally. Nor is this issue only of concern to professional ethicists. Kombe et al. (2019) find that some community members expressed skepticism regarding the use of randomization in a cluster RCT in Zambia on the grounds that access to the program under evaluation should be allocated to those who are most disadvantaged, not by means of a lottery.

On the one hand therefore, the ethical concern raised by randomization may only arise in a narrower set of cases than Karlan and Udry suppose, namely, cases where equipoise is not satisfied. On the other hand, the concern regarding fair allocation of a scarce good may also be more complicated than they suppose, not only requiring that there be no change in the expected aggregate value of the program, but also that participants in the RCT have equally strong claims to access it.

Counterfactual Policy

Karlan and Udry's (2020) appendix also ask researchers to consider the counterfactual policy: "Had the research not been conducted, is the counterfactual situation that would have happened instead predictably better for participants than what they actually received in any of the

arms of the study?" The concern here is for participants to not be subject to an inferior policy in the context of the RCT compared to the policy to which they would have been subject had the RCT not taken place. This question is important, but it might also be worth considering whether the ethically appropriate comparator is the policy to which participants *would otherwise* be subject – the counterfactual policy – or the policy to which participants *should* be subject.

As I discuss in a previous Transfer Project Blog Post (MacKay 2020b), this concern is addressed by the concept of "standard of care" in clinical research. The standard of care is the type and level of treatment that research participants should be guaranteed as a baseline. This concept is contested in clinical research ethics, but the interpretation of it as just *the care people can expect outside of the study* has largely been rejected in the clinical research ethics literature as too permissive. Applied to policy research, this interpretation would only prohibit assigning participants to interventions that are ex ante inferior to the counterfactual policy, regardless of how bad the status quo is. Indeed, as I discuss in my previous post, Aidan Corville, Sebastian Galiani, and Paul Gertler seem to adopt this position to defend their study with the Nairobi City Water and Sewerage Company which randomly assigned participants to a water disconnection intervention in cases of nonpayment. They write:

> The design of the study did not alter the service disconnection practices of Nairobi Water but rather affected which nonpaying households received the disconnection notices. Nairobi Water had always had a policy for service disconnection for nonpayment and was planning to scale this policy in the slums (Corville, Galiani, and Gertler 2020).

I've argued elsewhere that the relevant standard – at least for RCTs conducted or authorized by governments – should not be the counterfactual policy but rather the policy to which people have a claim of justice to be subject: researchers shouldn't randomize participants to interventions that are ex ante inferior to the intervention to which they *should* be subject (MacKay 2020a). Since this standard includes considerations of feasibility, Corville, Galiani, Gertler, and Yoshida's (2020) study may not run afoul of it, but it is worth considering whether "not making participants worse off" is too minimal a standard, and whether researchers should instead be asked to discuss the moral appropriateness of the interventions to which participants are randomly assigned. As with the sections on harms to participants, they should be asked to comment on this issue even if they are not responsible for the relevant decisions regarding implementation (for a more detailed discussion of research under non-ideal conditions, see MacKay 2020a).

An alternative way forward is to combine the scarcity and counterfactual policy sections of the structured ethics appendix into a "fairness of randomization" section. I argue elsewhere that the use of lotteries as a form of decision-making is valuable in cases of indeterminacy, that is, cases where the decision-maker has no sufficiently strong reason to prefer one option to another (MacKay 2020a). This means that randomization is permissible in two types of the cases. The first is where there is indeterminacy in the expert social scientific community regarding whether people should be assigned to either one or more arms of the study, or the intervention to which they are entitled. In this type of case, the social scientific community occupies a state of equipoise regarding which arm of the study best promotes the relevant interests of participants (MacKay 2020a). Importantly, this understanding of the concept of equipoise is more robust than the one endorsed by Karlan and Udry, and consistent with how the concept is employed in the clinical research setting.

The second type of case is where one arm of the study is expected to be superior in terms of realizing the relevant outcomes – e.g. outcomes regarding health, nutrition, or education – but the intervention/program is scarce. Here randomization is permissible if there is indeterminacy regarding *who* should receive access to the superior intervention (MacKay 2020a). This condition is satisfied when prospective participants or communities have equally strong claims to the scarce good, meaning allocation via lottery is permissible.

An alternative way forward therefore is to prompt researchers to answer the following questions:

> Is there equipoise regarding all arms of the study and the policy to which participants should have access? If not, was there a change in the expected aggregate value of the programs delivered? Do prospective participants have equally strong claims to the scarce programs?

This set of questions includes a robust concept of policy equipoise, which, I suggest above, is missing in the appendix, and also incorporates the ethical concerns Karlan and Udry express in their "scarcity" and "counterfactual policy" sections.

Community Engagement

Finally, one of the central ethical concerns that has been raised about policy research – particularly in development – is that it is designed and carried out by rich foreigners with insufficient concern for the priorities of local governments and communities (Deaton 2020; Ouma 2020). Clinical research ethicists have developed models of "community engagement," to address this concern, particularly for the context of international health research (Pratt and de Vries 2018). It is thus worth considering whether community engagement should be a binding norm for policy research, and what appropriate community engagement involves. In the context of Karlan and Udry's appendix, researchers could be given the opportunity to discuss their partnerships with local institutions and organizations to help address the above-mentioned concerns of critics.

Conclusion

Karlan and Udry's proposal for structured ethics appendices is welcome and timely, and their proposed appendix nicely captures the key ethical dilemmas researchers are likely to face in the design and conduct of their research. My critical comments are offered in a spirit of constructive engagement, and I recognize that they may require more than what it is reasonable to ask of researchers when it comes to the publication of their papers. More generally I would suggest that there is a real need for further systematic discussion of the ethics of policy research. Policy researchers need resources to identify and address the ethical dilemmas they face, and I fear research ethicists are currently playing catch-up.

References

Coville, Aidan, Sebastian Galiani, and Paul Gertler. 2020. "A Comment on the Ethical Issues Twitter Discussion in 'Enforcing Payment for Water and Sanitation Services in Nairobi's Slums.'" https://drive.google.com/file/d/1o0QibzPHUG1A_cTessYyeVn90429ejAQ/view.

Coville, Aidan, Sebastian Galiani, Paul Gertler, and Susuma Yoshida. 2020. "Enforcing Payment for Water and Sanitation Services in Nairobi's Slums," NBER Working Paper 27569.

Deaton, Angus. 2020. "Randomization in the Tropics Revisited: a Theme and Eleven Variations." NBER Working Paper 27600.

Emanuel, Ezekiel J., David Wendler, and Christine Grady. 2000. "What Makes Clinical Research Ethical?" *JAMA* 283: 2701-2711.

Freedman B. Equipoise and the ethics of clinical research. N Engl J Med. 1987 Jul 16;317(3):141-5. doi: 10.1056/NEJM198707163170304. PMID: 3600702.

Hoffman, Nimi. 2020. "Involuntary Experiments in Former Colonies: The Case for a Moratorium." *World Development* 127: 104805.

Karlan, Dean, and Christopher Udry. 2020. "A Call for Structured Ethics Appendices in Social Science Papers." Innovations for Poverty Action Blog, 12 November. https://www.poverty-action.org/blog/call-structured-ethics-appendices-social-science-papers

Kombe, Maureen Mupeta, Joseph Mumba Zulu, Charles Michelo, and Ingvild F. Sandøy. 2019. "Community Perspectives on Randomisation and Fairness in a Cluster Randomised Controlled Trial in Zambia." *BMC Medical Ethics* 20: 1-10.

London, Alex John. Forthcoming. "Equipoise: Integrating Social Value and Equal Respect in Research with Humans" In *The Oxford Handbook of Research Ethics*, edited by Ana Iltis and Douglas MacKay. https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780190947750.001.0001/oxfordhb-9780190947750-e-13

MacKay, Douglas. 2018. "The Ethics of Public Policy RCTs: The Principle of Policy Equipoise." *Bioethics* 32: 59-67.

MacKay, Douglas, and Averi Chakrabarti. 2019. "Government Policy Experiments and Informed Consent." *Public Health Ethics* 12: 188-201.

MacKay, Douglas. 2020a. "Government Policy Experiments and the Ethics of Randomization." *Philosophy and Public Affairs* 48: 319-352.

MacKay, Douglas. 2020b. "Enforcing Payment for Water Services – What's the Standard of Care?" The Transfer Project Blog, 14 August. https://transfer.cpc.unc.edu/wp-content/uploads/2020/08/Aug2020-Enforcing-payment-for-water-services-Doug-MacKay.pdf

Miller, Paul B., and Charles Weijer. 2003. "Rehabilitating Equipoise." *Kennedy Institute of Ethics Journal* 13: 93-118.

Ouma, Marion. 2020. "Trust, Legitimacy and Community Perceptions on Randomisation of Cash Transfers." Codesria, quarterly bulletin no.1. https://codesria.org/spip.php?article3085

Pratt, Bridget, and Jantina de Vries. 2018. "Community Engagement in Global Health Research that Advances Health Equity." *Bioethics* 32: 454-463.

Weijer, Charles, Jeremy M. Grimshaw, Martin P. Eccles, Andrew M. McRae, Angela White, Jamie C. Brehaut, Monica Taljaard, and the Ottawa Ethics of Cluster Randomized Trials Consensus Group. 2012. "The Ottawa Statement on the Ethical Design and Conduct of Cluster Randomized Trials." *PLoS Medicine* 9: 1-9.