What Role Should Equipoise Play in Experimental Development Economics?

Forthcoming in *Economics & Philosophy*

Penultimate draft, please cite published version when available

Marcos Picchio
Postdoctoral Fellow

Department of Bioethics, National Institutes of Health
10 Center Drive, Building 10, Bethesda, MD, 20892

Disclaimer: The views expressed in this article are my own and do not represent the position or policy of the NIH, DHHS, or US government.

Email: mpicchio@gmail.com URL: www.marcospicchio.com

Abstract: Unlike with randomized controlled trials (RCTs) in clinical research, little has been said about the ethical principles that should regulate the use of RCTs in experimental development economics. One well-known principle in clinical research ethics is the principle of clinical equipoise. Some recent commentators suggest that an analogue of clinical equipoise should play a role in experimental development economics. In this article, I first highlight some difficulties with importing the concept to experimental development economics. I then argue that MacKay's (2018, 2020) notion of *policy equipoise* avoids these difficulties and has a role to play in experimental development economics.

Keywords: development economics, field experiments, randomized controlled trials, clinical equipoise, research ethics

1. Introduction

In 2007, Cohen and Dupas (2010) conducted a field experiment on the effects of subsidies on preventive healthcare products in low-income countries. To conduct this experiment, Cohen and Dupas selected twenty prenatal clinics in Kenya and varied the price at which they could sell insecticide-treated bed nets (ITNs) to pregnant women. ITNs are known to prevent malaria infection and consequently are also highly effective in reducing maternal anemia and infant mortality in regions where malaria is prevalent. The widespread use of ITNs also generates positive externalities in the form of health benefits to non-users. In the experiment, sixteen of the clinics were randomly assigned to four different groupings corresponding to different subsidy levels ranging from a full subsidy (free) to a 90% subsidy. The remaining four clinics were used as a control and did not have access to the highly subsidized bed nets.

The experiment described above is a notable example of what Banerjee and Duflo (2009) have dubbed the "experimental approach to development economics" (henceforth, experimental development economics) (for overviews see Banerjee and Duflo 2011; Karlan and Appel 2011). A central feature of experimental development economics is the use of randomized controlled trials (RCTs) to learn about the effects of socioeconomic interventions (e.g. cash transfers or product subsidies) on questions relating to economic development. To date, economists commenting on the use of RCTs in development economics have largely focused on methodological issues pertaining to the scientific value of such experiments (e.g. Banerjee and Duflo 2009; Kremer and Holla 2009; Rodrik 2009; Barrett and Carter 2010; Deaton 2010; Harrison 2011; Basu 2014; Deaton and Cartwright 2018; Heckman 2020; Ravallion 2020).[1] Significantly less attention, however, has been paid to ethical questions surrounding the design and implementation of RCTs in development economics (henceforth, development RCTs).[2] This is somewhat surprising since the use of RCTs in clinical research is subject to intense ethical scrutiny by physicians and bioethicists alike.

[1] Also see Boone and Johnson (2009: 61-64) and Faverau (2016) for discussion of some methodological differences between clinical trials and development RCTs.

[2] The earliest and most systematic discussion is Baele (2013). Some early contributions from economists include Alderman *et al.* (2016), Glennerster and Powers (2016), Ziliak and Teather-Posadas (2016). More recent contributions include Abramowicz and Szafarz (2020), Hoffmann (2020), Drèze (2023), Khera (2023).

To many outside observers, the experiment described at the outset seems like a clear candidate for such ethical scrutiny. Yet Cohen and Dupas do not explicitly address ethical concerns anywhere in the article in which they published their findings, which reflects a general trend in experimental development economics.[3] Some may take the lack of direct engagement with ethical questions as a sign that development economists have borrowed well-established experimental methods from the biomedical sciences and ignored the stringent ethical requirements that accompany their use. In this article, I take steps towards assessing the accuracy of this claim by focusing on the role that the concept of *equipoise* should play in experimental development economics.

In its most general guise, equipoise is a state of uncertainty about the relative merits of a set of interventions (London 2024). Many physicians and bioethicists believe that some form of equipoise is a necessary component of ethical clinical trial design. The most well-known formulation of this requirement (to be discussed in more detail below) is the *principle of clinical equipoise* (Freedman 1987). The principle states that, for a clinical trial to proceed in an ethically defensible manner, the expert medical community needs to be in a state of collective uncertainty with respect to the relative therapeutic benefits of each trial arm.

To date, what little has been said about the ethics of development RCTs has centered on equipoise. Commentators such as Baele (2013) and Abramowicz and Szafarz (2020) suggest that development economists should care about equipoise.[4] These commentators appear optimistic that equipoise can help development economists think through difficult ethical questions. Other commentators, such as Ziliak and Teather-Posadas (2016), go a step a further and criticize development economists for routinely failing to comply with the principle of clinical equipoise. As this article reveals, these suggestions and criticisms are misguided.

Despite the misgivings alluded to above, equipoise does have a role to play in experimental development economics. Specifically, MacKay's (2018, 2020) notion of *policy equipoise* can play the role of justifying the fairness of randomization in some experimental contexts. Establishing this conclusion requires careful analysis and argument. What is crucial for understanding the position I advance is a proper understanding of the two roles that equipoise plays in clinical

[3] Asiedu *et al.* (2021) have recently taken steps towards reversing this trend by calling for all articles in experimental development economics to contain an ethics appendix.
[4] Deaton (2010) and Ravallion (2020) also make similar suggestions in passing.

research: (1) resolution of an ethical tension between the role obligations of physicians and the scientific enterprise; and (2) providing evidence that research is socially valuable. Both these roles are analyzed and discussed in what follows.

An outline of the rest of this article is as follows. Section 2 provides a brief history of equipoise for readers unfamiliar with the concept. In sections 3-5, I highlight various difficulties with importing equipoise to experimental development economics. In doing so, I propose desiderata that a principle of equipoise should meet for the concept to be relevant to experimental development economics. Section 6 shows how policy equipoise meets my proposed desiderata and how it is relevant to experimental development economics. Section 6 also makes the abstract discussion of policy equipoise concrete by analyzing the experiment discussed at the outset. Section 7 discusses an important complication with how policy equipoise applies experimental development economics. Section 8 concludes.

2. A Brief History of Equipoise

A major theme in the field of research ethics is the tension between advancing scientific knowledge and respecting the rights of human subjects. These rights impose correlative obligations on researchers. For example, all humans possess a right to bodily integrity, and this imposes a correlative obligation on others to obtain consent for a bodily intrusion. The obligation to respect bodily integrity applies regardless of whether one is a physician, economist, or anyone, for that matter. But not all obligations are like this. Some obligations are *special* in that they are owed to a subset of persons. Notably, physicians have special obligations to their patients in virtue of voluntarily occupying an institutionally specified social role (Hardimon 1994). One of the most well-recognized ethical obligations physicians owe their patients is the *therapeutic obligation*:

> *Therapeutic Obligation*: A physician should recommend a treatment *T* for some condition *C* if and only if *T* is accepted by the medical community as effective in treating *C*.[5]

Patients, in turn, have a correlative right that their physician comply with this obligation.

The physician's therapeutic obligation is key to understanding the initial motivation behind the development of equipoise as an ethical requirement for clinical trials. As the use of RCTs in clinical research became more widespread, there was growing concern that the new experimental

[5] Some commentators will refer to this ethical requirement as the duty of personal care.

methodology was in tension with the therapeutic obligations of physicians (Hellman and Hellman 1991). Throughout, I refer to this (alleged) tension as the *RCT Dilemma:*

> *RCT Dilemma*: If a physician *P* has reason to believe that new Therapy A is better than another Therapy B, then *P* cannot permissibly enroll or advise patients to enroll in a clinical trial of A versus B because, ethically, *P* is obligated to recommend A to each new patient with a need for one of these therapies.[6]

Early commentators worried that physicians conducting clinical trials routinely face this dilemma because physicians often develop "treatment preferences" (Schafer 1982) based on incomplete scientific evidence, personal experience, and mere hunches. If a physician preferring Therapy A to Therapy B enrolls patients in a trial comparing A to B, the physician would then (supposedly) violate her therapeutic obligation and the clinical trial would therefore be unethical. Instead of being a *genuine* dilemma where two equally bad options are presented, what the RCT Dilemma really is then, is a particular instance of the general tension between advancing scientific knowledge and respecting the rights of human subjects.[7]

The concept of equipoise was first introduced by Fried (1974) as a way of resolving the RCT Dilemma.[8] Fried reasoned that the only way a physician could permissibly enroll patients into a clinical trial without violating her therapeutic obligation was if she was genuinely uncertain about the relative therapeutic benefits of each trial arm. The distinguishing feature of Fried's version of equipoise—or *theoretical* equipoise—is that the uncertainty necessary for permissible randomization had to be in the mind of the individual physician. Because Fried did not elaborate on what he meant by genuine uncertainty, commentators interpreted him as suggesting that physicians enrolling patients in a clinical trial need to maintain that the probability of Therapy A being superior to Therapy B is exactly 50 percent (Chalmers 1978; cf. Miller and Weijer 2003). This made equipoise seem like an unlikely state to be in given the tendency for physicians to form treatment preferences. Even if a physician found herself in such a state due to the lack of any available evidence, critics complained that equipoise was fragile and would easily be disturbed as soon as a clinical trial commenced and evidence favoring one trial arm accrued (Marquis 1983;

[6] The first commentators to make note of this dilemma appear to be Shaw and Chalmers (1970: 487), from which my formulation of the RCT Dilemma is drawn.
[7] My choice of terminology follows Miller and Joffe (2011).
[8] Equipoise was not the only solution proposed. See Marquis (1983), Gifford (1986), and Freedman (1987) for discussions of other proposed solutions to the RCT Dilemma.

Gifford 1986; Freedman 1987). This posed a problem because clinical trials would need to be prematurely stopped before any statistically significant results could be established.[9]

The difficulties with Fried's version of equipoise led Freedman (1987) to propose clinical equipoise as a necessary requirement of ethical clinical trial design. Per London (2024), Freedman's clinical equipoise is meant to solve the same ethical problem as theoretical equipoise (cf. Miller and Weijer 2003). However, clinical equipoise differs from theoretical equipoise in that it requires uncertainty to obtain in the expert medical community rather than the mind of the individual physician. Interestingly, one of the central motivations for this shift is an appeal to the social dimensions of scientific knowledge (Longino 1990). Freedman (1987: 144) writes that medicine "is social rather than individual in nature" and its advancement relies on "progressive consensus within the medical and research communities." Crucially, this means that for clinical equipoise to obtain, the individual members of the relevant medical community do *not* each need to be in a state of theoretical equipoise (though this is certainly a possibility). Instead, for there to be genuine uncertainty about the relative therapeutic benefits of Therapy A to Therapy B, there can exist "an honest, professional disagreement among expert clinicians about the preferred treatment" (Freedman 1987: 144). Because establishing clinical equipoise involves disagreement rather than indifference, it allows physicians to have treatment preferences and allows for a clinical trial to proceed until sufficient evidence is accumulated to resolve uncertainty in the expert medical community.

Having specified the relevant sense of uncertainty, we can turn to Freedman's influential formulation of the principle of clinical equipoise:

> *Principle of Clinical Equipoise*: "at the start of a trial, (1) there must be a state of clinical equipoise regarding the merits of the regimens to be tested, and (2) the trial must be designed in such a way as to make it reasonable to expect that, if it is successfully concluded, clinical equipoise will be disturbed" (Freedman 1987: 144; numbers added).

Note that I have separated the principle of clinical equipoise into two components. This is done to emphasize the following: while there is no lack of professional disagreement among economists (i.e. a "state of equipoise" with respect to the merits of some policy intervention), economics is

[9] One practical solution to this problem that has now become standard practice in clinical research is to conceal interim analyses from investigators, research participants, and funders, and only allow an independent Data Monitoring Committee having access to interim results. This Data Monitoring Committee will be contractually obligated to withhold interim results unless certain statistical thresholds are met.

unlike biomedical science in that there is sometimes a lack of consensus on how such disagreements can be resolved (i.e. how to "disturb equipoise"). This is a point I expand on below.

3. Methodological Considerations

Any form of equipoise should incorporate something like the second component of Freedman's formulation of the principle of clinical equipoise. This should be uncontroversial. You cannot resolve any collective uncertainty if nothing new is learned. This leads to the first desideratum that a principle of equipoise should meet if it is to play a role in experimental development economics:

> Desideratum #1: A principle of equipoise should require that a development RCT be designed in such a way as to make it reasonable to expect that, if it is successfully concluded, equipoise will be disturbed (i.e. something will be learned).

As plausible as this first desideratum is, there are complications with it. As the rest of this section makes clear, more needs to be said about how it is that a development RCT can be designed in such a way as to make it reasonable to expect that, if it is successfully concluded, equipoise will be disturbed.

RCTs are often regarded as the "gold standard" when it comes to evaluating the effectiveness of some clinical intervention (cf. Cartwright 2007). While one clinical trial may not provide enough evidence to completely resolve all (reasonable) professional disagreement, there is agreement in the medical community that a series of clinical trials *is* often the best way to resolve the professional disagreement in question. This is another way of saying that the results of a well-designed clinical trial can disrupt the medical community's state of equipoise with respect to the merits of some biomedical intervention. Freedman tacitly assumed this methodological consensus when devising the principle of clinical equipoise. The principle of clinical equipoise would thereby deem a statistically underpowered clinical trial as unethical since it would do little to disrupt the medical community's state of equipoise with respect to a disputed intervention. Though this is intuitive enough, it is not clear how this important aspect of the principle of clinical equipoise derives from the therapeutic obligations of physicians—a point I return to in section 5.

Unlike in biomedical science, there is much less consensus among economists about the scientific value of RCTs in their discipline (see anthologies such as Cohen and Easterly 2009a;

Ogden 2017; Bédécarrats *et al*. 2020a). The absence of methodological consensus poses a difficulty for those who think that equipoise should play a role in experimental development economics: How can an RCT resolve any sort of professional disagreement if not all members of the relevant professional community believe that RCTs *can resolve* professional disagreements in the first place? It would be hasty to simply ignore this problem by dismissing the lack of methodological consensus as another instance of the reluctance economists have historically had towards experimentation (see Guala 2005: 2-3). As the discussion below reveals, the worries about the scientific value of RCTs in development economics are worth taking seriously.

For evidence of the disagreement about the scientific value of RCTs, consider the well-known methodological criticisms raised by Angus Deaton.[10] For present purposes, I will restrict my attention to Deaton's (2010) discussion of the problem of external validity (or generalizability).[11] While an RCT does—in principle—provide an unbiased estimate of the mean causal effect of an intervention, matters are not so simple in the social scientific context. Skeptics often argue that an RCT is a "black box" method of causal inference: "A treatment is administered, an outcome is observed, with no need for any understanding of what is going on in between and *why* a treatment produces its outcome" (Reiss 2013: 205). The criticism in question stems from the fact that an RCT alone does not provide much in the way of a scientific explanation. Specifying the relevant sense of "scientific explanation" is important. When it comes to the social sciences, a leading contender is causal mechanistic explanation (Elster 2015). The lack of explanatory evidence generated by an RCT significantly contributes to the problem of external validity since drawing inferences beyond the immediate test population typically requires knowledge of how the treatment produced the outcome of interest. Deaton (2010) notably argues that evidence from development RCTs will be of little use in furthering our understanding of economic development, and thereby making successful interventions, if efforts are not taken to also identify the casual mechanisms that generate the effects uncovered through field experimentation (also see Acemoglu

[10] See Ogden (2020) for a survey of the various critiques of RCTs in development economics and a discussion of how development economists conducting RCTs have changed their research practices to address these critiques.
[11] Of course, internal validity matters as well. There is little point in worrying about external validity if internal validity is compromised. And as Deaton and Cartwright (2018) suggest, the distinction between internal and external validity may ultimately be misleading. I focus on external validity because it has been at the center of methodological debates surrounding development RCTs and for good reason: discovering generalizable knowledge is an oft-cited goal of scientific research.

2010). More concerning is that policymakers can misapply findings from an RCT if they do not take the problems such as external validity seriously (Cartwright and Hardie 2012).

It is worth acknowledging that external validity is also a problem for clinical trials, though the problem is often ameliorated by a corpus of background biomedical knowledge (Backmann 2017). While a corpus of undisputed social scientific knowledge is harder to come by (Reiss 2019), Deaton suggests that the problem of external validity in development RCTs can at least be addressed by closer engagement with economic theory:

> It is certainly not always obvious how to combine theory with experiments. Indeed, much of the interest in RCTs…comes from a deep skepticism of economic theory, and impatience with its ability to deliver structures that seem at all helpful in interpreting reality. Applied and theoretical economists seem to be further apart now than at any period in the last quarter century. Yet failure to reintegrate is hardly an option because without it there is no chance of long-term scientific progress or of maintaining and extending the results of experimentation. RCTs that are not theoretically guided are unlikely to have more than local validity, a warning that applies equally to nonexperimental work (Deaton 2010: 450).

In addition to closer engagement with economic theory, Deaton (2010) also suggests that experimental design should incorporate previous empirical findings, particularly from behavioral economics. On Deaton's view, experiments guided by theory and previous empirical findings are more likely to generate the knowledge of causal mechanisms necessary to make successful socioeconomic interventions. Going forward, I will take these two considerations to be necessary (but not sufficient) components of sound experimental design.

While I have no illusions that methodological disputes in economics can be easily settled with appeals to ethical considerations, there is an important takeaway when it comes to the role equipoise should play in experimental development economics: to satisfy desideratum #1, experimental design needs to consider both economic theory and previous empirical findings (among other things). This is so that the results from an experiment have some chance of "disturbing equipoise" in the expert development community.

4. Equipoise and the RCT Dilemma in Experimental Development Economics

This section responds to an obvious difficulty with importing the most familiar version of equipoise to experimental development economics: any straightforward application of the principle of clinical equipoise to experimental development economics would render the great

majority of development RCTs unethical if the principle is considered a necessary condition for permissible experimentation. As noted at the outset, Ziliak and Teather-Posadas (2016) have harped on this very point to criticize development RCTs. As this section reveals, this criticism is misguided because economists do not face an analogue of the RCT Dilemma. The main takeaway from this section is that the first role of equipoise (resolution of the RCT Dilemma) is not straightforwardly applicable to experimental development economics.

Unlike with a new treatment for a serious medical condition, it is true that there is not much mystery surrounding the therapeutic benefits of an insecticide-treated bed net (e.g. Cohen and Dupas 2010) or a deworming tablet (e.g. Miguel and Kremer 2004). The point generalizes to interventions that are not healthcare related such as cash transfers and access to other valuable resources and opportunities. We can claim these things are typically good for people without appealing to a substantive theory of well-being (Hausman 2011), much less a scientific experiment. If this is right, then to abide by a direct analogue of the principle of clinical equipoise, development economists would have to maintain—with a straight face—that there is no way of knowing whether some obviously beneficial intervention really is—in fact—welfare enhancing without first conducting an RCT.[12] This would undeniably put development economists in an extremely awkward position.

To avoid unnecessary confusion, it is important that development economists conducting RCTs make clear that the goal of an experiment is not to determine that an intervention provides benefits to individual recipients (though this is a foreseeable consequence). Rather, for a large class of development RCTs, the goal of an experiment is to determine *how* effective an intervention is *for* achieving a specific socioeconomic outcome *in* a population of interest. All outcomes of interest may ultimately be grounded in a concern with human welfare, but this is not the same as saying that "we just do not know" whether giving someone a valuable resource is good for them.[13]

[12] It is important to acknowledge that appealing to intuition or common sense is not always a reliable way of knowing whether some intervention is beneficial. There are plenty examples throughout the history of medicine of interventions thought to be "obviously beneficial" that turned out to be harmful. But this point can be easily abused and used to reach absurd conclusions, e.g. an RCT on the effectiveness of parachutes (Smith and Pell 2003). See Basu (2014) for more on this point as applied to development RCTs.

[13] One may still wonder why the goal of an experiment stated above cannot serve as the basis for a principle of equipoise in experimental development economics. For example, on London's (2001) account of equipoise in international human-subjects research, clinical equipoise is ultimately concerned with the net therapeutic benefit of a set of interventions relativized to a particular health problem in a specific treatment setting. Kukla's (2007) principle of equipoise, which I discuss further on, incorporates this kind of consideration in a way that is relevant to determining the social value of an experiment.

For example, we know that deworming drugs cure a parasitic infection that makes people unhealthy, but what we really may be in interested in is whether this drug increases school attendance. This kind of knowledge is socially valuable because it helps determine whether an intervention *should* be made accessible to a population (Kukla 2007).

An objection to the argument above is that it only shows that development economists need to elevate their ethical standards to the level found in clinical research. If development RCTs routinely violate clinical equipoise, then it is development economists who need to adapt—not a well-known ethical requirement. But this objection overlooks a more serious difficulty: economists do not stand in any sort of professional relation to research subjects and therefore do not have special obligations to provide "care" (i.e. access to socioeconomic resources and opportunities) in virtue of the social role they occupy.[14] Because development economists lack the relevant role obligations, they do not face an analogue of the RCT Dilemma, which provides the motivation for the concept of equipoise in the first place.[15] Picchio (2024) delves into the issue further and concludes that, barring emergency scenarios, development economists do not have *any* obligations to provide research subjects with "care". This is important because, without an analogue of the physician's therapeutic obligation for development economists, there is once again no straightforward basis for an RCT Dilemma and therefore no need for equipoise.

[14] Development economists still have negative duties to not harm research participants (or anyone, for that matter) in virtue of being persons, not researchers. But it is also important to stress that strict adherence to the injunction "do not harm" is untenable as documents such as *The Belmont Report* make clear: learning what is harmful or beneficial sometimes requires exposing research participants to some risks (National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research 1978). Risk-benefit assessment is the primary means by which Institutional Review Boards (IRBs) operationalize the "do not harm" principle. However, clinical equipoise is orthogonal to these risk-benefit assessments for reasons that Miller (2012) makes clear. For a helpful discussion of risk-benefit assessment in experimental development economics (which makes no reference to equipoise), see Glennerster and Powers (2016).

[15] It is worth mentioning that Miller and Brody (2003, 2007), two of the most notable critics of clinical equipoise, argue that equipoise is irrelevant to the ethics of clinical research precisely because they see the RCT Dilemma as a pseudo-problem to begin with. For these critics, the ends of clinical practice differ from the ends of clinical research. The former is concerned with providing individualized treatment to patients and the latter with discovering generalizable medical knowledge for the benefit of society. As a result, physicians conducting clinical research (arguably) do not stand in a therapeutic relationship to research participants. Consequently, the ethical principles regulating clinical practice (e.g. the requirement that physicians provide the best-known treatment for some medical condition) are inappropriate in the context of clinical research, which should be governed by a separate set of ethical principles. While assessing this argument is beyond the scope of this article (see Chiong 2006), it is important to keep in mind that, contrary to what some commentators suggest, there is no consensus among bioethicists regarding the role equipoise should play in the ethics of clinical research. See Chiong (2006), Veatch (2007), and Miller and Joffe (2011) for additional critiques of clinical equipoise.

The second desiderata that a principle of equipoise in experimental development economics should accommodate is as follows:

Desideratum #2: A principle of equipoise should account for an analogue of the RCT Dilemma in experimental development economics.

5. Social Equipoise

The most well-known justification of clinical equipoise is that it resolves the RCT Dilemma. This section discusses another role that equipoise plays in clinical research and which does not make any reference to the therapeutic obligations of physicians. The role I have in mind is that of providing evidence that research is socially valuable. To distinguish the two roles, I call the version of equipoise discussed in this section *social equipoise*.[16]

5.1 Equipoise as Evidence of Social Value

In a widely cited article on the ethical requirements of clinical research, Emanuel *et al*. (2000) identify *value* as the first requirement of ethical clinical research. These commentators maintain, quite plausibly, that clinical research is valuable when it is directed at an intervention that could lead to improvements in health or well-being. This is not to suggest that value is the *only* ethical requirement, but rather that it would be difficult to justify human subjects research without *some* reference to the concept. This is why the second ethical requirement Emanuel *et al*. (2000) identify is scientific validity since, for clinical research to be valuable, it needs to be conducted in a methodologically rigorous manner. What is noteworthy about Emanuel *et al*.'s (2000) framework is that it does not rely on the therapeutic obligations of physicians. Instead, the primary ethical reason given for why clinical research needs to be valuable and scientifically valid is that finite resources should be used responsibly.[17] What I take Emanuel *et al.* (2000) to mean by "responsible use of finite resources" is that researchers should be sensitive to ethical opportunity costs. Considering that estimates for conducting a development RCT range anywhere between $500,000

[16] The term "social equipoise" is not (to my knowledge) commonly used. I borrow the term from Petticrew *et al.*'s (2013) call for such a concept to be developed.

[17] Emanuel *et al*. (2000) also cite avoidance of exploitation as a secondary reason. Exploitation is a notoriously difficult concept to analyze and on which to get clarity, so I avoid discussing it here. See Hawkins and Emanuel (2008) for an anthology devoted to the topic of exploitation in international clinical research. The issue of exploitation is one that a complete account of the ethics of development RCTs will have to address.

to $1,500,000 (Bédécarrats *et al.* 2020b: 24; cf. Khera 2023: 6), it should be uncontroversial to suggest that this ethical reason also carries over to research in experimental development economics.

Despite not relying on the therapeutic obligations of physicians, Emanuel *et al*. (2000) do suggest a different role for the principle of clinical equipoise. Instead of providing a solution to the RCT Dilemma, clinical equipoise can play the role of indicating that research is scientifically valid and in turn valuable. As Emanuel *et al*. (2000: 2704) argue, "If there exists a consensus about what is the better treatment, there is no null hypothesis, and the research is invalid." With respect to value, they add that "without clinical equipoise, research that compares therapies is unlikely to be of value because the research will not contribute to increasing knowledge about the best therapy…" (2000: 2704). Note here that these considerations make good on the promise of a clearer explanation for why a poorly designed clinical trial is unethical. The statistically underpowered trial is unethical because it lacks scientific validity, and therefore, fails to use finite resources in a responsible manner.

What I take Emanuel *et al*. (2000) to be suggesting is that clinical equipoise plays an *evidentiary* role in clinical research. Below, I explore whether this conception of equipoise (which I am calling *social equipoise*) can play a similar role in experimental development economics. To be more precise, I explore whether social equipoise can play the role of indicating when an experiment has *social value* (which for present purposes I distinguish from *clinical value*).

Because development economists (and their funders) have obligations to use finite resources responsibly and can discharge such obligations by conducting (and sponsoring) socially valuable research (Pierson and Millum 2018), it is important to explore whether the concept of social equipoise can play a role in ensuring that these obligations are complied with. As made clear below, social equipoise does not refer to any special obligations that researchers owe research participants, nor does it require researchers to awkwardly maintain agnosticism with respect to the therapeutic value or individual welfare benefits of some interventions. Instead, the role social equipoise can play is that of indicating when a development RCT is socially valuable. However, as I argue below, social equipoise is either trivial or superfluous when applied to experimental development economics.

5.2 Kukla's Equipoise

The connection between equipoise and social value has, to date, been most thoroughly explored by Kukla (2007). As Kukla makes clear, "My version of the principle of equipoise does not focus on equipoise with respect to the relative expected outcomes of trial arms, but rather on equipoise concerning the social value of the intervention being tested" (2007: 173). Though Kukla does not use the term "social equipoise," their version of the principle of equipoise is a notable first attempt at articulating such a principle.

> *Kukla's Principle of Equipoise*: "In order to begin or continue human subjects research, one must be in a state of equipoise with respect to whether or the extent to which the intervention being tested *should* be made accessible to the population that falls under the scope of the research" (Kukla 2007: 180).

An important feature of Kukla's principle is that it provides helpful criteria for determining when human subjects research is socially valuable. Another important feature of Kukla's principle of equipoise is that it is *not* grounded in the therapeutic obligations of physicians nor any special obligations for that matter. Instead, Kukla's account is grounded in general ethical requirements of justice and respect for persons, which I assume can generate an obligation to use finite resources responsibly. However, Kukla's principle is not without complications. Below, I respond to these complications by proposing two additional desiderata that any principle of equipoise should meet if it is to apply to experimental development economics.

The first two complications concern the locus of uncertainty for social equipoise. Because Kukla seems to assume that researchers qua persons have duties to provide others with what they are owed as a matter of justice, there is undue concern with equipoise obtaining in the mind of the individual researcher, thus indicating that all researchers face an analogue of the RCT Dilemma with respect to their duties to provide others what they are due. There are two issues with this. The first issue is that Kukla assumes that individuals are required to provide others with what they are owed as a matter of justice. As MacKay (2015) argues, it is the responsibility of social institutions and not individual researchers (which includes economists) to provide research subjects with what they are owed as matter of justice (also see Picchio 2024).

The second issue is that Kukla's principle of equipoise suffers from the same fatal deficiency as theoretical equipoise, namely, that it is incredibly fragile. It seems unlikely that an individual development economist can ever be genuinely uncertain as to whether some intervention should be made accessible to a population of interest. This is especially true if

development economists are engaging in sound experimental design by consulting economic theory and previous empirical findings. Consequently, one can expect development economists to form judgments (akin to treatment preferences) about whether some intervention they are investigating should be made accessible. This is especially true once an experiment commences.

If the locus of uncertainty for equipoise is not in the mind of the individual researcher, then it should be in the relevant expert community. Like with clinical equipoise, there can be honest, professional disagreement in the expert development community. But this honest, professional disagreement should not be confined to just the experimental subset of the expert development community. To fully acknowledge the social dimensions of scientific inquiry as the principle of clinical equipoise does, development economists need to design their experiments in a way that can result in something being learned; this means attempting to design experiments in a way that could convince non-experimental colleagues that the outcome of an RCT could have some bearing on a research question of interest. As already stressed above, experimental design should incorporate both economic theory and previous empirical findings so that the results from the experiment have some chance of "disturbing social equipoise" in the wider expert development community.[18]

Combining these two considerations we arrive at a third desideratum:

Desideratum #3: A principle of equipoise should specify that the locus of uncertainty is the expert development community.

The final complication concerns the necessary and sufficient conditions for establishing that an experiment is socially valuable. Kukla claims that social equipoise requires that there be uncertainty about whether an intervention should be made accessible to a population of interest. Though this may be sufficient for establishing social value, it is certainly not necessary. There are multiple ways in which a development RCT can be socially valuable. In addition to using RCTs for impact evaluation, Banerjee and Duflo (2009) also promote using RCTs to estimate key parameters in economic models used for policymaking—a presumably socially valuable research activity. Taking cue from Deaton and Cartwright's (2018) discussion of the scientific value of

[18] Specifying the relevant criteria for membership into the expert development community is a challenge that I do not explore here. This is a difficulty that the principle of clinical equipoise also faces (see Veatch 2007).

RCTs, Morduch (2020) has recently emphasized that there are two distinct ways in which RCTs have been used by development economists. The first use of RCTs is for the familiar purpose of evaluating the impact of socioeconomic interventions. The use of "what works" RCTs (as Morduch calls them) for impact evaluation corresponds well with Kukla's criteria for establishing social value.[19] But Morduch also identifies a more promising use of RCTs in development economics; these are "how and why" RCTs. "How and why" RCTs can be used to answer research questions that are "exploratory, theory-driven, and motivated by the desire to understand economic possibilities and constraints" (Morduch 2020: 109). As such, "how and why" RCTs can be used to gain valuable insights into the nature of economic contracts, behaviors, and institutions.

"How and why" RCTs often make use of short-term limited-scale programs and do not necessarily have bearing on whether the treatment being administered should be made accessible to a wider population.[20] Still, what is learned through such experiments could be socially valuable. Using an RCT to confirm the prediction of a theory or provide a counterexample to a general theoretical proposition could have important downstream consequences (Deaton and Cartwright 2018). If equipoise is to be relevant to experimental development economics, the concept should not rule out a "how and why" RCT in which the treatment arm receives a highly beneficial intervention that cannot be made accessible to the population from which the sample is drawn. For example, consider an experiment that, for the purposes of testing economic theory, briefly provides the treatment arm with very generous unconditional cash transfers that are not sustainable long term (e.g. Egger *et al*. 2022). Such an experiment could still prove socially valuable and be permissible to conduct so long as additional ethical criteria are met.[21]

The final desideratum that a principle of equipoise should meet is as follows:

Desideratum #4: A principle of equipoise should recognize the different ways a development RCT can be socially valuable.

[19] It is also worth highlighting that "what works" RCTs are the development RCTs that have been the central focus of the methodological criticisms discussed earlier.
[20] See Morduch (2020) for analysis of actual examples of "how and why" RCTs.
[21] Some criteria discussed by Emanuel *et al.* (2000) and MacKay (2024) relevant here include favorable risk-benefit ratio and fair subject selection.

5.3 Formulation of the Principle of Social Equipoise

At this stage, it may be tempting to accommodate all desiderata mentioned so far by proposing a principle such as the following:

> *Principle of Social Equipoise:* At the start of an experiment, (1) there must be disagreement or doubt (i.e. equipoise) in the expert development community with respect to whether or the extent to which the intervention being tested should be made accessible to a population of interest or has bearing on a research question of interest and (2) the experiment must be designed in such a way as to make it reasonable to expect that, if it is successfully concluded, something will be learned: there will be some effect on the extent of disagreement or doubt (i.e. equipoise will be disturbed).

Careful inspect reveals that the principle above accommodates desiderata #1-4. But a complaint about the principle that I have formulated above is that it is not much of an ethical principle but rather a statement of sound research practice. In fact, some will be quick to point out that, when clarified, equipoise simply amounts to a more familiar concept: scientific relevance. If this is correct, one would have to wonder why it would be even helpful to talk in terms of equipoise at all. The principle above could be seen as specifying what the concept of "scientific relevance" entails. But if equipoise simply amounted to a statement of sound research practice, then it would be safe to conclude that the concept has no (substantive) role to play in the ethics of experimental development economics. However, there is a more recent version of equipoise to consider, which avoids the complications I have identified, and which I turn to now.

6. Policy Equipoise

In a recently proposed framework for incorporating ethical considerations into experimental development economics, Asiedu *et al*. (2021) suggest that MacKay's (2018, 2020) notion of *policy equipoise* should play a role in experimental development economics. In this section, I make a positive case for this suggestion by arguing that policy equipoise meets the four desiderata I have identified. I also clarify some misconceptions about policy equipoise and begin assessing the question of how it applies to experimental development economics, which I continue to explore in section 7.

6.1. What is Policy Equipoise?

Douglas MacKay has developed the concept of policy equipoise in response to the increasing use of RCTs for public policy research (henceforth, policy RCTs) by local, state, or federal

governments and their respective research agencies. This will already strike some as a complication since not all development RCTs appear to qualify as policy RCTs. After all, development RCTs are often conducted in collaboration with non-public entities such as privately funded non-governmental organizations (NGOs) and private for-profit corporations (PFPCs) such as microfinance banks (e.g. Banerjee *et al*. 2015) and utility companies (e.g. Devoto *et al*. 2012). Though private actors are not ethical freelancers, it is uncontroversial to maintain that private entities have different ethical duties than governments. For ease of exposition, I set this complication aside and turn back to it in section 7.

In MacKay's (2020) framework for permissible randomization, the physician's therapeutic obligation is replaced by a government's duty to promote justice-related outcomes for its citizens. MacKay argues that for any justice-related outcome a government has a duty to realize, governments possess an obligation to implement the policy that is "(1) evidence-based, (2) consistent with people's rights, and (3) consistent with the realization of other target outcomes" (MacKay 2020: 323). MacKay calls this policy the Best Proven, morally and practically Attainable and sustainable (BPA) policy. A policy is "morally and practically attainable and sustainable" if and only if "(1) it is consistent with residents' rights and (2) it can be implemented for an appropriate period of time given a just system of resource procurement and allocation" (MacKay 2020: 324).

Because governments are required to implement BPA policies, governments and their respective research agencies are also required to treat citizens fairly by not subjecting anyone to policies known to be inferior to the BPA policy. This generates an analogue of the RCT Dilemma for government-authorized researchers conducting a policy RCT, and MacKay's notion of policy equipoise is *one* possible way to justify the fairness of randomization in policy research contexts. But since policy research often takes place in conditions in which BPA policies are not in place, MacKay (2020) in fact proposes two principles of policy equipoise. I distinguish the two principles by whether they apply to ideal vs. non-ideal circumstances:

> *Principle of Policy Equipoise (Ideal)*: "Government agencies may randomly assign participants to different policy interventions if they are in a state of genuine equipoise regarding all arms of the study and the BPA policy" (MacKay 2020: 329).
>
> *Principle of Policy Equipoise (Non-Ideal)*: "A government agency may randomly assign participants to different policy interventions if: (1) it occupies a state of genuine equipoise

> regarding all arms of the study and the non-BPA status quo policy and (2) it does not have the authority to implement the BPA policy" (MacKay 2020: 333).

Given the circumstances of the developing world (resource limitations, state capacity, etc.), some will be inclined to think that the first policy equipoise principle will never apply to experimental development economics. However, one could argue that since the notion of a BPA policy includes feasibility constraints, it factors in considerations such as resource limitations and state capacity into the definition of the best proven and *attainable* policy.

Whether it is the ideal or the non-ideal version of policy equipoise that matters requires determining if BPA policies are currently in place in development context.[22] This question detracts from the more important question at hand, which is, whether policy equipoise is relevant to experimental development economics in the first place. Below, I use the desiderata I identified in sections 3-5 to tackle this question and to answer it affirmatively.

6.2. Applying the Desiderata

I begin with the fourth desideratum to address what careful observers may see as a limitation with policy equipoise, which is, in fact, a positive feature. Recall that the clinical equipoise is regarded as a *necessary* condition for permissible randomization. Though MacKay (2018) initially argued that policy equipoise is a necessary condition for permissible randomization, MacKay's (2020: 321) considered view is that policy equipoise is a *sufficient* condition for permissible randomization. As already mentioned, this means that policy equipoise is *one* way of justifying the fairness of randomization in a policy RCT.

At first glance, it is not obvious how this consideration ties into the fourth desideratum, which calls for a principle of equipoise to acknowledge the different ways a development RCT can be socially valuable. Making policy equipoise a sufficient condition does just this because, as MacKay (2020) and MacKay and Cohn (2023) discuss in more detail, there are circumstances where governments have good reasons to experiment without being in a state of policy equipoise. For example, a government may want to evaluate a BPA Policy against an intervention known to be superior in realizing target outcomes, but which is not practically attainable or sustainable. A government may want to do this for estimating cost-effectiveness or to gain a better understanding

[22] Picchio (2024) argues that BPA policies are rarely in place in the development context, so it is the non-ideal version of policy equipoise that is relevant to experimental development economics.

of the causal mechanism which generates the outcome of interest. MacKay's (2020) framework provides criteria for when randomization is permissible in scenarios like these without making any reference to policy equipoise. What this means is that policy equipoise does not automatically deem a government-sponsored "how and why" RCTs as unethical. It further means that policy equipoise will not be a relevant ethical consideration in every experimental context. Making policy equipoise a sufficient condition limits its applications to experimental development economics. However, this is a feature of policy equipoise, not a bug.

Desiderata #2 and #3 are more straightforwardly met. Policy equipoise accounts for an analogue of the RCT Dilemma that government-authorized researchers conducting a policy RCT may potentially face. With respect to desiderata #3, MacKay (2020: 328) suggests that it is the relevant social scientific community that is in equipoise. Presumably, this social scientific community includes policy experts in positions of power—not just academic researchers. I will illustrate this point with the Cohen and Dupas (2010) experiment discussed at the outset, but first say something about the methodological considerations captured by desideratum #1.

Though MacKay (2020) does not address methodological considerations, it is easy to supplement his two policy equipoise principles so that they meet this desideratum. What is required is the addition of a second component to each principle that states the following: "the trial must be designed in such a way as to make it reasonable to expect that, if it is successfully concluded, policy equipoise will be disturbed." This simple modification prevents researchers from easily and conveniently citing expert disagreement as justification for an experiment that is unlikely to "disturb equipoise". Additionally, this modification ensures that policy equipoise meets the four desiderata I have identified in this article.

6.3. *Cohen and Dupas Revisited*

The foregoing discussion of equipoise has been abstract. To complete the positive case for policy equipoise, I analyze Cohen and Dupas (2010) experiment through the lens of policy equipoise to illustrate the role it should play in experimental development economics. By reviewing the experiment, my aim is to illustrate how policy equipoise applies to experimental development economics and how it can be invoked to justify the fairness of randomization.

Hearing the details of Cohen and Dupas's (2010) experiment undoubtedly makes outside observers uneasy. Why not provide free ITNs to as many people as possible if they are known to

be therapeutically effective against malaria and known to generate positive externalities? As Cohen and Dupas make clear at the start of their article, prior to their experiment there was well-known professional disagreement over whether health products which generate positive externalities should be given away for free or whether a positive subsidized price should be charged. The debate was undoubtedly a question about which intervention promoted a development-related outcome (reducing malaria transmission) more effectively. What is noteworthy is that this debate had two prominent development economists advocating for conflicting positions. Sachs (2005) was in favor of free distribution and relied on standard economic theory to argue that making ITNs free would lead to higher uptake and use. Easterly (2006) opposed free distribution and cited empirical findings from behavioral economics to suggest that fully subsidizing ITNs can have the opposite effect. Easterly reasoned that charging a positive price could induce people to use ITNs due to psychological sunk-cost effects (Thaler 1980). Also noteworthy is that this debate had policy ramifications, though crucially neither of the agencies involved were government entities (at least, initially). The World Health Organization (WHO)—an agency of the United Nations—had sided with Sachs while Population Services International—a nonprofit global health organization—had sided with Easterly.[23]

Prior to experimentation, it was clear that there was disagreement in the expert development community over the BPA policy with respect to full versus partial subsidization of ITNs. What's more, since this disagreement was based on a clash between standard economic theory and previous empirical findings, a well-designed RCT could generate evidence relevant to the debate. This is not to suggest that evidence from one experiment could definitively settle the debate, as some may have initially been inclined to think.[24] Though Cohen and Dupas do not use the term "policy equipoise" anywhere in their article, these researchers provide evidence that the expert development community was in a state of policy equipoise with respect to the issue of partial versus full subsidization of ITNs.

Though the experiment was privately funded, the fact that there was policy equipoise was important because Cohen and Dupas worked closely with the Kenyan Ministry of Health to

[23] See Kremer and Holla (2009) for an overview of the more general debate over how pricing affects take-up of education and health services and products, and how evidence RCTs has contributed to this policy debate.
[24] See Rodrik (2009: 28-39) for a more detailed analysis and commentary. As Rodrik makes clear, Cohen and Dupas's experiment is not immune from external validity critiques. Cf. Kremer and Holla (2009).

conduct their experiment.[25] Cohen and Dupas (2010: 11) report that there were 70 health clinics in the region they were interested in studying. These 70 health clinics were a mixture of public government run health clinics and private NGO-run health clinics. One of the criteria which Cohen and Dupas used to narrow their sample was public status. Ultimately, Cohen and Dupas only included 20 health centers in their sample. The 20 health clinics that were randomized into different subsidy levels were all public, i.e. managed and operated by the Kenyan Ministry of Health.

Cohen and Dupas (2010) ultimately found that free distribution of ITNs led to both higher uptake and usage than partial subsidization (or no subsidization whatsoever). What's more, these findings have played a key role in subsequent policy decisions by NGOs operating in low-income countries as well as the policy decisions of governments.[26] But what is important to note is that the Kenyan women affected by the experiment were not treated unfairly by the researchers or their respective government (another basis for ethical complaint is discussed in closing). And this is because there was policy equipoise, i.e. there was a state of uncertainty over whether full subsidization really was the right course of policy and whether pregnant women were entitled to free ITNs in the first place (i.e. it was unclear what the BPA policy was). None of this should be taken to suggest that randomizing a policy intervention is always fair or all things considered permissible (MacKay and Cohn 2023), but only that Cohen and Dupas (2010) provides a helpful example of how policy equipoise applies to experimental development economics.

7. Does Policy Equipoise Extend Beyond Government Policy Experiments?

The previous section argued that policy equipoise has a role to play in some experimental contexts. There remains the question of policy equipoise's scope since, on first appearances, not all development RCTs qualify as policy RCTs (and vice versa—policy RCTs are conducted in high-income countries too). In this section, I outline two strategies for extending policy equipoise beyond *government* policy RCTs. While I reject the first strategy, I argue that the second strategy

[25] The acknowledgements section for Cohen and Dupas (2010) states: "We thank the Mulago Foundation for its financial support, and the donors to TAMTAM Africa for providing free nets distributed in this study. Jessica Cohen was funded by a National Science Foundation Graduate Research Fellowship. We are grateful to the Kenyan Ministry of Health and its staff for their collaboration."

[26] For example, GiveWell directed $100 million to free ITN distribution after Cohen and Dupas's experiment (Ogden 2020: 127). Cohen and Easterly (2009b: 18) also report that the experiment played a key role in the Kenyan government's eventual decision to fully subsidize ITNs for pregnant women.

is more promising. Ultimately, I suggest that more research needs to be done into the special obligations of NGOs and other private actors conducting experiments.

7.1. The Public vs. Private Distinction

The clearest case in which policy equipoise can be invoked to justify the fairness of randomization is in the case of an "embedded experiment" (Drèze 2023), where development economists work closely with a government agency to evaluate a policy. As seen above, Cohen and Dupas (2010) provides one helpful illustration of how policy equipoise applies in such contexts. In these contexts, development economists act as government-authorized researchers and are thereby incur special obligations that they would otherwise not have. This means justifying the fairness of randomization to research subjects (and a host of additional ethical considerations that space does not permit a full exploration of here).

Matters are less straightforward when it comes to development RCTs that are sponsored or conducted in collaboration with an NGO or some other private entity. There is an interesting question of the extent to which NGOs operating in developing countries share the same obligations as governments. NGOs in developing countries do perform some of the functions of governments, most notably, public good provision (Werker and Ahmed 2008). But it would be puzzling to suggest that an organization incurs the same ethical obligations as a government simply because it supplies a public good. Philanthropists supply public goods, yet we do not expect the same degree of ethical accountability from philanthropists as we do from governments.[27] The relevant difference between philanthropists and governments is that governments rely on coercion to supply public goods. And it is the coercive nature of government that grounds the special obligations it owes its citizens (Blake 2001; Nagel 2005). In the context of discussing policy equipoise, we have focused on the obligation governments have to treat citizens fairly. But NGOs and other private actors lack the requisite coercive capacities (which is not to deny that private actors have some coercive capacities). The question then remains of what NGOs owe the people they serve, which is important for assessing the ethical permissibility of experimentation. However, there are two strategies for sidestepping the issue, which I evaluate below.

[27] Perhaps we *should* demand more accountability from philanthropists, but this is not the same as holding philanthropists to the same ethical standard as governments. See Reich (2018) on this issue.

7.2. The Case for Extending Policy Equipoise Beyond Government Policy Experiments

There is an important way in which one could make the case that an NGO-led development RCT can fall under the scope of policy equipoise. This is through public support for an experiment. Though private entities may supply most of the funding for a development RCT, many development RCTs are large projects whose total funding also comes from public sources. For example, it is not uncommon for the first round of follow-up measurements to be funded by the NGO or business implementing the program at stake, with further follow-up funded by standard research grants.[28] More generally, as long as researchers from publicly funded universities (or from private universities receiving public funding) are conducting RCTs, then there is a sense in which public support is being deployed for the realization of the experiment.

What public funding for (most, if not all) development RCTs suggests is that development economists incur the same obligations as government-authorized researchers *even if* it is an NGO or PFPC that is implementing the experimental intervention. If development economists incur the same obligations as government-authorized researchers, then they would also have the same obligations to treat research subjects fairly. Policy equipoise would thereby become available as a potential justification for the fairness of randomization. This proposal is in some ways not too far from how ethical oversight for human subjects research is justified in countries such as the United States. All human subjects research conducted or *funded* by the U.S. federal government is subject to ethical regulation (Common Rule 2018). This notably includes human subjects research conducted by university-affiliated researchers since (both public and private) universities receive public support for their activities.

The case for extending policy equipoise via public support has promise but requires modification. While I do not deny that receiving public support through funding sources should come with the caveat that research undergo ethical review, this is different from maintaining that researchers receiving public resources incur the same ethical obligations as governments. Maintaining such a position would severely limit what kind of research is conducted since researchers would have to adopt the same justice-based research priorities as governments. Given the ubiquity of public funding for research, this would not only restrict individual freedom, but it would also stifle innovation. More generally, if receiving public funding were sufficient for

[28] I am grateful to anonymous referee for pointing this out to me and raising the possibility that policy equipoise could apply in cases where there is public support for an experiment.

incurring the ethical responsibilities of governments, then taken to its logical extreme, this proposal would effectively eliminate the private sphere all together.

Despite the complications raised above, it is well worth noting that public support is not just limited to funding sources. This is especially important for the case of embedded experiments. Even if entirely privately funded, conducting a development RCT typically requires communication and collaboration with government actors, which means that scarce human resources are diverted away from other government activities.[29] Cohen and Dupas (2010) is once again a good example of this. The acknowledgements section of the article suggests the experiment was privately funded, but we saw above how the experiment required cooperation and coordination with the Kenyan Ministry of Health. While private actors conducting experiments do not have the same obligations as governments, the employees of government agencies (acting in their capacities as government employees) should still act in ways consistent with the special obligations of the government employing them. So should the higher-level government actors that may authorize the use of scarce human resources for the realization of an experiment. This means, among other things, not administering or authorizing interventions that treat citizens unfairly. Just like how one may maintain that physicians conducting clinical trials are still bound by their therapeutic obligation, the point is that assisting with an experiment does not suddenly absolve government actors of the ethical obligations they have qua government actors.

7.3. Taking Stock

To sum up, I have argued that receiving public funding is not sufficient for extending policy equipoise beyond the realm of government policy RCTs. However, I have argued that, if a government's scarce human resources are employed for the realization of an experiment, then there is a case for extending policy equipoise beyond government policy experiments. This is because the government actors (acting in their capacity as government actors) involved in the experiment must act consistently with their obligations to treat citizens fairly. This is the case even if an experiment is privately funded.

There also remains the possibility, which I have not explored here, that NGOs and other private actors have substantive special obligations to those they interact with that go beyond a negative duty to do no harm. Getting clear on the nature of these special obligations would shed

[29] London (2005: 30) discusses this consideration in the context of international clinical research.

light on the ethics of development RCTs. For example, if NGOs have duties of justice like that of governments, then NGOs would have obligations to treat those they serve fairly, and policy equipoise would become available as a potential justification for the fairness of randomization. However, the antecedent of this conditional is bound to be controversial and requires careful investigation, which cannot be undertaken here.

8. Conclusion

In closing, it is crucial to emphasize again that policy equipoise is *not* a sufficient condition for permissible experimentation. In other words, policy equipoise is not a "green light" to proceed with an experiment. To reiterate, policy equipoise has a specific role to play in experimental development economics, namely, that of justifying the fairness of randomization in certain experimental contexts (such as a "what works" RCT). This means that there will be other experimental contexts (such as a "how and why" RCT) in which it would be inappropriate to ask whether policy equipoise obtains.

To drive the point home, note that above I only argued that the Kenyan women involved in Cohen and Dupas's (2010) experiment were not treated unfairly when they were randomized into different trial arms. This is different from concluding that there are no other ways in which these women may have been wronged. Like many development RCTs, Cohen and Dupas's (2010) experiment employed a cluster randomized design, which often makes obtaining individual consent for some interventions impracticable.[30] Of pressing importance is analyzing how risks to research subjects can be justifiably imposed given that many development RCTs involve cluster randomization.[31] Even in situations where some form of consent can be obtained, there is also the troubling possibility of imposing risks on bystanders due to the tendency for socioeconomic interventions to have unintended spillover effects. These issues should be high ethical priorities for development economists conducting experiments and tools from economic analysis could prove helpful here. While the vast literature on clinical research ethics contains some valuable

[30] In Cohen and Dupas's (2010) experiment, women did consent to have their hemoglobin levels recorded for data collection purposes. But none of these women consented to having the price of ITNs manipulated by researchers.
[31] The lack of informed consent in development RCTs has been documented and criticized by Hoffmann (2020). MacKay and Chakrabarti (2019) provide some helpful criteria for judging when policy experimentation without consent is permissible, but more work remains to be done on this front.

lessons for development economists (see MacKay 2024; Picchio 2024), development economists can and should also reflect more on the ethical tools at their disposal.

Acknowledgements

I would like to thank Robert Steel and two anonymous referees for comments and suggestions that helped improved this article. I would also like to thank Dan Hausman, Norm Fost, Paul Kelleher, Doug MacKay and participants at the Second Lake Como Summer School on Economic Behaviours for helpful feedback on earlier versions. This research was supported (in part) by the Intramural Research Program of the National Institutes of Health. The views expressed in this article are my own and do not represent the position or policy of the NIH, DHHS, or US government.

References

Abramowicz, M. and A. Szafarz. 2020. Ethics of RCTs: Should Economists Care about Equipoise? In *Randomized Control Trials in the Field of Development: A Critical Perspective*, (ed.) F. Bédécarrats, I. Guérin and F. Roubaud, 280-292. New York: Oxford University Press.

Acemoglu, D. 2010. Theory, General Equilibrium, and Political Economy in Development Economics. *Journal of Economic Perspectives*, 24(3): 17-32.

Alderman, H., J. Das, and V. Rao. 2016. Conducting Ethical Economic Research: Complications from the Field. In *Oxford Handbook on Professional Economic Ethics*, ed. G. DeMartino and D. McCloskey, 402-422. New York: Oxford University Press.

Asiedu, E., D. Karlan, M.P. Lambon-Quayefio, and C.R. Udry. 2021. A Call for Structured Ethics Appendices in Social Science Papers. *PNAS*, 118(29).

Baele, S. 2013. The Ethics of New Development Economics: Is the Experimental Approach to Development Economics Morally Wrong? *The Journal of Philosophical Economics*, 7(1): 2-42.

Backmann, M. 2017. What's in a Gold Standard? In Defence of Randomized Controlled Trials. *Medicine, Health Care, and Philosophy*, 20: 513-523.

Banerjee, A. and E. Duflo. 2009. The Experimental Approach to Development Economics. *The Annual Review of Economics* 1(1): 151-178.

Banerjee, A. and E Duflo. 2011. *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*. New York: PublicAffairs.

Banerjee, A., E. Duflo, R. Glennerster, and C. Kinnan. 2015. The Miracle of Microfinance? Evidence from a Randomized Evaluation. *American Economic Journal: Applied Economics,* 7(1): 22-53.

Barrett, C. and M. Carter. 2010. The Power and Pitfalls of Experiments in Development Economics: Some Non-Random Reflections. *Applied Economic Perspectives and Policy*, 32(4): 515-548.

Basu, K. 2014. Randomisation, Causality and the Role of Reasoned Intuition. *Oxford Development Studies*, 42(4): 455-472.

Bédécarrats, F., I. Guérin and F. Roubaud, eds. 2020a. *Randomized Control Trials in the Field of Development: A Critical Perspective*. New York: Oxford University Press.

Bédécarrats, F., Guérin, I., and Roubaud, F., 2020b. Editor's Introduction: Controversies Around RCT in Development In *Randomized Control Trials in the Field of Development: A Critical Perspective*, ed. F. Bédécarrats, I. Guérin and F. Roubaud, 1-28. New York: Oxford University Press.

Blake, M. 2001. Distributive Justice, State Coercion, Autonomy. *Philosophy & Public Affairs* 30(3): 257-296.

Boone, P. and S. Johnson. 2009. Breaking out of the Pocket: Do Health Interventions Work? Which ones and in What Sense? In *What Works in Development: Thinking Big and Thinking Small*, ed. J. Cohen and W. Easterly, 55-83. Washington, DC: Brookings Institute Press.

Cartwright, N. 2007. Are RCTs the Gold Standard? *BioSocieties* 2(1): 11-20.

Cartwright, N. and J. Hardie. 2012. *Evidence-Based Policy: A practical guide to doing it better*. New York: Oxford University Press.

Chalmers, T.C. 1978. The Ethics of Randomization as a Decision-Making Technique and the Problems of Informed Consent. In *Contemporary Issues in Bioethics*, ed. T. Beauchamp and L. Walters, 426-249. Encino, CA: Dickenson.

Chiong, W. 2006. The Real Problem with Equipoise. *American Journal of Bioethics*, 6(4): 37-47.

Cohen, J. and P. Dupas. 2010. Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment, *The Quarterly Journal of Economics,* 125(1): 1-45.

Cohen, J., and W. Easterly, eds. 2009a. *What Works in Development: Thinking Big and Thinking Small*. Washington, DC: Brookings Institution Press.

Cohen, J. and W. Easterly. 2009b. Introduction: Thinking Big Versus Thinking Small. In *What Works in Development: Thinking Big and Thinking Small*, ed. J. Cohen and W. Easterly, 1-23. Washington, DC: Brookings Institute Press.

Common Rule, 45 CFR 46. 2018. https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/index.html.

Deaton, A. 2010. Instruments, Randomization, and Learning About Development. *Journal of Economic Literature* 48(2): 424–455.

Deaton, A. and N. Cartwright. 2018. Understanding and Misunderstanding Randomized Controlled Trials. *Social Science & Medicine* 210: 2-21.

Devoto, F., E. Duflo, P. Dupas, W. Parienté and V. Pons. 2012. Happiness on Tap: Piped Water Adoption in Urban Morocco. *American Economic Journal: Economic Policy*, 4(4): 68-99.

Drèze, J. 2023. The perils of embedded experiments. *Review of Development Economics*: 1-13.

Easterly, W. 2006. *The White Man's Burden: Why the West's Efforts to Aid the Rest Have Done So Much Ill and So Little Good*. New York: Penguin Press.

Egger, D., J. Haushofer, E. Miguel, P. Niehaus and M. Walker. 2022. General equilibrium effects of cash transfers: experimental evidence from Kenya. *Econometrica*, 90(6): 2603-2643.

Elster, J. 2015. *Explaining Social Behavior: More Nuts and Bolts for the Social Sciences*. Cambridge: Cambridge University Press.

Emanuel, E., D. Wendler, D. and C. Grady. 2000. What Makes Clinical Research Ethical? *JAMA*, 283(20): 2701-2711.

Freedman, B. 1987. Equipoise and the Ethics of Clinical Research. *The New England Journal of Medicine*, 317(3): 141-145.

Fried, C. 1974. *Medical Experimentation: Personal Integrity and Social Policy*. New York: Elsevier.

Favereau, J. 2016. On the analogy between field experiments in economics and clinical trials in medicine. *Journal of Economic Methodology* 23(2): 203-222.

Glennerster, R. and S. Powers. 2016. Balancing Risk and Benefit: Ethical Tradeoffs in Running Randomized Evaluations. In *Oxford Handbook on Professional Economic Ethics*, ed. G. DeMartino and D. McCloskey, 367-401. New York: Oxford University Press.

Gifford, F. 1986. The Conflict Between Randomized Clinical trials and the Therapeutic Obligation. *The Journal of Medicine and Philosophy*, 11(4): 347-366.

Guala, F. 2005. *The Methodology of Experimental Economics*. Cambridge: Cambridge University Press.

Hardimon, M.O. 1994. Role Obligations. *Journal of Philosophy* 91(7): 333-363.

Harrison, G. 2011. Randomisation and its Discontents. *Journal of African Economies*, 20(4): 626-652.

Hausman, D. 2011. *Preference, Value, Choice, and Welfare*. Cambridge: Cambridge University Press.

Hawkins, J. and E. Emanuel, eds. 2008. *Exploitation and Developing Countries: The Ethics of Clinical Research*. Princeton: Princeton University Press.

Heckman, J. 2020. Epilogue: Randomization and Social Policy Evaluation Revisited. In *Randomized Control Trials in the Field of Development: A Critical Perspective*, ed. F. Bédécarrats, I. Guérin, and F. Roubaud, 304-330. New York: Oxford University Press.

Hellman, S. and D.S. Hellman. 1991. Of Mice but Not Men: Problems of the Randomized Clinical Trial. *New England Journal of Medicine* 324 (22): 1585-1589.

Hoffmann, N. 2020. Involuntary Experiments in former colonies: The case for a moratorium. *World Development*, 127: 1-3.

Karlan, D. and J. Appel. 2011. *More than Good Intentions: Improving the Ways the World's Poor Borrow, Save, Farm, Learn, and Stay Healthy*. New York: Penguin.

Khera, R. 2023. Some questions of ethics in randomized controlled trials. *Review of Development Economics*: 1-16.

Kremer, M. and A. Holla. 2009. Pricing and Access: Lessons from Randomized Evaluations in Education and Health. In *What Works in Development: Thinking Big and Thinking Small*, ed. J. Cohen and W. Easterly, 91-119. Washington, DC: Brookings Institute Press.

Kukla, R. 2007. Resituating the Principle of Equipoise: Justice and Access to Care in Non-Ideal Conditions *Kennedy Institute of Ethics Journal*, 17(3): 171-202.

London, A.J. 2001. Equipoise and International Human-Subjects Research. *Bioethics* 15(4): 312-332.

London, A.J. 2005. Justice and the Human Development Approach to International Research. *Hastings Center Report*, 35(1): 24-37.

London, A.J. 2024. Equipoise: Integrating Social Value and Equal Respect in Research with Humans. In *The Oxford Handbook of Research Ethics*, ed. D. MacKay and A. Iltis, 216-235. Oxford: Oxford University Press.

Longino, H.E. 1990. *Science as Social Knowledge*. Princeton: Princeton University Press.

MacKay, D. 2015. Standard of Care, Institutional Obligations, and Distributive Justice. *Bioethics* 29(4): 262-273.

MacKay, D. 2018. The Ethics of Public Policy RCTs: The Principle of Policy Equipoise. *Bioethics* 32(1): 59-67.

MacKay, D. 2020. Government Policy Experiments and the Ethics of Randomization. *Philosophy & Public Affairs*, 48(4): 319-352.

MacKay, D. 2024. The Ethics of Public Policy Experiments: Lessons from Clinical Research Ethics. In *The Oxford Handbook of Research Ethics*, ed. D. MacKay and A. Iltis, 445-472. Oxford: Oxford University Press.

MacKay, D. and A. Chakrabarti. 2019. Government Policy Experiments and Informed Consent. *Public Health Ethics*, 12(2): 188-201.

MacKay, D. and E. Cohn. 2023. Public Policy Experiments Without Equipoise: When is Randomization Fair? *Ethics & Human Research*, 45(1): 15-28.

Marquis, D. 1983. Leaving Therapy to Chance. *The Hastings Center Report* 13(4): 40-47.

Miguel, E. and M. Kremer. 2004. Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities. *Econometrica*, 72(1): 159-217.

Miller, F.G. 2012. Clinical Equipoise and Risk-Benefit Assessment. *Clinical Trials*, 9: 621-627.

Miller, F.G. and H. Brody. 2003. A Critique of Clinical Equipoise. *The Hastings Center Report* 33(3): 19-28.

Miller, F.G. and H. Brody, H. 2007. Clinical Equipoise and the Incoherence of Research Ethics. *Journal of Medicine and Philosophy*, 32(2): 151-165.

Miller, F.G. and S. Joffe. 2011. Equipoise and the Dilemma of Randomized Clinical Trials. *The New England Journal of Medicine*, 364(5): 476-480.

Miller, P.B. and C. Weijer: 2003. Rehabilitating Equipoise. *Kennedy Institute of Ethics Journal*, 13(2): 93-118.

Morduch, J. 2020. The Disruptive Power of RCTs. In *Randomized Control Trials in the Field of Development: A Critical Perspective*, ed. F. Bédécarrats, I. Guérin and F. Roubaud, 108-125. New York: Oxford University Press.

Nagel, T. 2005. The Problem of Global Justice. *Philosophy & Public Affairs* 33(2): 113-147.

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. 1978. *The Belmont Report: ethical principles and guidelines for the protection of human subjects of research.* Washington, DC: US Department of Health, Education, and Welfare.

Ogden, T., ed. 2017. *Experimental Conversations: Perspectives on Randomized Trials in Development Economics*. Cambridge, MA: MIT Press.

Ogden, T. 2020. RCTs in Development Economics, Their Critics and Their Evolution. In *Randomized Control Trials in the Field of Development: A Critical Perspective*, ed. F. Bédécarrats, I. Guérin and F. Roubaud.126-151. New York: Oxford University Press.

Petticrew, M., M. McKee, K. Lock, J. Green and G. Phillips. 2013. In Search of Social Equipoise. *BMJ*, 347.

Picchio, M. 2024. What is the standard of care in experimental development economics? *Politics, Philosophy & Economics*, 23(2): 205-226.

Pierson, L. and J. Millum. 2018. Health Research Priority Setting: The Duties of Individual Funders. *American Journal of Bioethics*, 18(11): 6-17.

Pritchett, L. 2020. Randomizing Development: Method or Madness? In *Randomized Control Trials in the Field of Development: A Critical Perspective*, ed. F. Bédécarrats, I. Guérin and F. Roubaud, 79-107. New York: Oxford University Press. New York: Oxford University Press.

Ravallion, M. 2020. Should the Randomistas (Continue to) Rule? In *Randomized Control Trials in the Field of Development: A Critical Perspective*, ed. F. Bédécarrats, I. Guérin and F. Roubaud, 49-78. New York: Oxford University Press. New York: Oxford University Press.

Reich, R. 2018. *Just Giving: How Philanthropy is Failing Democracy and How it Can Do Better*. Princeton: Princeton University Press.

Reiss, J. 2013. *Philosophy of Economics: A Contemporary Introduction*. New York: Routledge.

Reiss, J. 2019. Expertise, Agreement, and the Nature of Social Scientific Facts or: Against Epistocracy. *Social Epistemology*, 33(2): 183-192.

Sachs, J. 2005. *The End of Poverty: Economic Possibilities for Our Time*. New York: Penguin Press.

Schafer, A. 1982. The Ethics of the Randomized Clinical Trial. *The New England Journal of Medicine*, 307(12): 719-724.

Shaw, L.W. and T.C. Chalmers. 1970. Ethics in Cooperative Clinical Trials. *Annals of the New York Academy of Sciences*, 169(2): 487-495.

Smith, G. C. and J.P. Pell. 2003. Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. *BMJ*, 327(7429): 1459-1461.

Thaler, R. 1980. Toward a Positive Theory of Consumer Choice. *Journal of Economic Behavior & Organization*, 1(1): 39-60.

Veatch, R.M. 2007. The Irrelevance of Equipoise. *The Journal of Medicine and Philosophy*, 32(2): 167-183.

Werker, E. and F.Z. Ahmed. 2008. What do Nongovernmental Organizations Do? *Journal of Economic Perspectives*, 22(2): 73-92.

Ziliak, S. and E. Teather-Posadas. 2016. The Unprincipled Randomization Principle in Economics and Medicine. In *Oxford Handbook on Professional Economic Ethics*, ed. G. DeMartino and D. McCloskey, 423-452. New York: Oxford University Press.

Biographical Information

Marcos Picchio is currently a postdoctoral fellow in the Department of Bioethics at the National Institutes of Health. He completed a PhD in philosophy at the University of Wisconsin-Madison in Spring 2023.