Journal of Economic Behavior & Organization 123 (2016) 31–56

Contents lists available at ScienceDirect

Journal of Economic Behavior & Organization

journal homepage: www.elsevier.com/locate/jebo

Partner selection into policy relevant field experiments☆

Michèle Belot[a], Jonathan James[b,*]

[a] University of Edinburgh, United Kingdom
[b] University of Bath, United Kingdom

ARTICLE INFO

Article history:
Received 7 April 2015
Received in revised form
17 December 2015
Accepted 18 December 2015
Available online 30 December 2015

JEL classification:
C93
I18
J13

Keywords:
Selection
Field experiments
Randomised controlled trials
External validity

ABSTRACT

This study investigates the issue of self-selection of stakeholders into participation and collaboration in policy-relevant experiments. We document and test the implications of self-selection in the context of randomised policy experiment we conducted in primary schools in the UK. The main questions we ask are (1) is there evidence of selection on key observable characteristics likely to matter for the outcome of interest and (2) to what extent does selection matter. The experimental work consists in testing the effects of an intervention aimed at encouraging children to make more healthy choices at lunch. We recruited schools through local authorities and randomised schools across two incentive treatments and a control group. We document the selection-taking place both at the level of local authorities and at the school level. Overall we find mild evidence of selection on key observables such as obesity levels and socio-economic characteristics. We find evidence of selection along indicators of involvement in healthy lifestyle programmes at the school level, but the magnitude is small.

1. Introduction

Field experiments in economics and the social sciences have become increasing popular (Holt, 2005). The main driving factors behind this increasing prevalence are, on the one hand, the quest for identification of causal mechanisms – which is easier to achieve when researchers are directly involved in manipulating the economic environment of interest – and, on the other hand, a quest to remain close to reality as opposed to studying subjects in an isolated laboratory context. There is now a stronghold of researchers advocating the case for randomized controlled trials (RCTs) in social policy (Burtless, 1995; Duflo and Kremer, 2005).

This study is interested specifically in a fundamental methodological issue associated with field experimental research: the selection of field collaborators into the experiment. Conducting field experiments usually requires finding collaborators such as employers, policymakers, schools, etc. who are prepared to collaborate with researchers and provide the necessary support for data collection. As List (2011) puts it, the support of a key person prepared to stand behind the research project is often critical: "*Have a champion within the organization – the higher up the better. Making the experiment a "we" project instead*

☆ We thank the Esmee Fairbairn Foundation for their financial support (grant number 10-2206). We thank seminar participants at the University of Bath, Oxford, Stirling, and conference participants at the European Economic Association conference (Malaga, 2012), Royal Economic Society conference (Royal Holloway, London, 2013), Florence Workshop on Behavioural and Experimental Economics (May 2013), International Conference on Public Policy (Grenoble, June 2013).

* Corresponding author. Tel.: +44 07786260013.
E-mail address: j.james@bath.ac.uk (J. James).

http://dx.doi.org/10.1016/j.jebo.2015.12.007
0167-2681/

of an "us versus them" pursuit as early as possible is critical." Surprisingly perhaps, field experimental papers devote very little attention to the issue of "selection into the experiment". As an illustration, we provide in Table A1 a brief overview of the information provided in field experimental studies published in the top 5 journals and in the American Economic Journal: Applied Economics. We focus on the fields of policy evaluation, personnel economics and development economics, which have all experienced a significant increase in the popularity of field experimental research.[1] In most cases we know little or nothing about how the collaborators were selected and approached, and the experimental sample is not compared to the broader population of interest. One notable exception is a recent paper by Fryer (2011).

Researchers are well aware of the limitations that result from restricting experiments to subjects who have opted in. The main limitation is a possible selection bias and a lack of external validity. Of course, some form of selection is inevitable. An experiment will, for example, take place within a given geographical area and at a particular point in time. This initial selection is often for practical reasons. Researchers located in California will find it more practical to conduct a field experiment in California than across the entire United States or across several countries. It is probably even desirable that the experiment can be conducted with sufficient oversight by the researchers. It does nevertheless raise issues of generalizability, insofar as it restricts the sample to a population with certain characteristics (e.g. people living in California at a particular time).

How important is selection in field experimental research? Harrison and List (2004) acknowledge in their review paper that we know very little about the implications of self-selection for field experimental research. Seven years later Ludwig et al. (2011) point out that this is still an open question that has not been answered. At the moment, researchers tend to be very conservative and modest in their claims regarding the external validity of their results. We believe that a proper documentation of the selection process would help us drawing more general lessons from field randomised controlled experiments.

We study self-selection in participation in an experiment conducted in a highly policy relevant domain: children's diet. The goal of the experiment is to test the effectiveness of various incentive schemes to encourage children to eat fruit and vegetables at lunch.[2] To conduct this experiment, we sought the collaboration of primary schools in England. We recruited them through local education authorities (LEAs), which play an overarching and coordinating role. Rather than picking a set of local authorities in an arbitrary manner, we approached all local authorities in the country at the same time and in the same manner (via e-mail) and asked whether they would be interested in collaborating with us. If they responded positively, we asked them to bring us in contact with at least five local schools representative of the local authority. Providing names of schools requires some effort, so the type of selection we study is not only based on initial interest (that is low cost to indicate) but also on actual commitment in the experiment. The randomisation eventually takes place at the school level and within local authorities, so local authorities should expect some schools to be treated and some schools to be part of a control group. The schools are the ones that are ultimately directly involved in the experiment and data collection. We contacted the schools suggested by the local authorities, briefed them about the project and they then decided whether to participate or not. Thus, we have potential selection operating at different levels: self-selection of local authorities, selection of "representative" schools by local authorities and selection of schools into the experiment.

We document how selection operates at these different levels along observable characteristics of the population under consideration – characteristics that we would expect could matter in the decision to participate, such as obesity rates and socio-economic indicators. We consider a wide range of variables that could a priori be relevant and see whether they are correlated with selection or not. Then we investigate whether the treatment effects are biased by observables correlated with selection, we do not find any evidence that this is the case.

It is clear that we cannot control for some key variables (such as personality characteristics of the people involved) that may introduce a selection bias in the RCT as well. Moreover, it is clear that if selection was only driven by observables, it would be straightforward to correct for it. The point here is to get a sense of how much selection takes place along characteristics that could a priori be relevant and are observable, which should in principle give a sense of the importance of the full selection problem.

Our findings can be summarised as follows. First, we find that out of the 150 local authorities we initially contacted, only 12 eventually participated in the actual experiment. We find that local authorities who express initial interest tend to be larger and richer, and have less favourable characteristics in terms of the outcome of interest (e.g. lower rates of fruit and vegetable consumption), but the selection is mild. We find no evidence that selection operates according to pre-trends. We also find little evidence of selection at the school level, except for one dimension, which is that the schools suggested by the local authorities to conduct the experiment are more likely to be involved in programmes promoting healthy lifestyles (as evaluated by an independent official body). Second, we do not find any significant correlations between the treatment effects of the experiment and the variables which, albeit to a mild degree, are correlated with selection into the experiment.

Of course, the selection we document here is also "case-specific". We cannot claim that the selection we document informs us about the magnitude and type of selection taking place in other policy-relevant field experiments. But in the domain of policy-relevant field experiments, this is an example of an intervention that targets a "hot" topic on the policy agenda (children's obesity) and the experimental intervention we propose is typical in the sense that we propose to compensate the partners for the costs involved with conducting the intervention and collecting the data. Even in that case we find that

[1] We searched these journals systematically for the keywords 'field experiment'.
[2] We refer to Belot et al. (2015) for the full analysis of the experiment.