

PREFERENCES FOR INCOME DISTRIBUTION AND DISTRIBUTIVE JUSTICE:

A WINDOW ON THE PROBLEMS OF USING EXPERIMENTAL DATA IN ECONOMICS AND ETHICS

Norman Frohlich
University of Manitoba

and

Joe Oppenheimer
University of Maryland

In ethics, as in other fields struggling to find an empirical basis, most knowledge claims are initially based on, and justified by common sense, every-day observations and deep introspection.¹ Theoretical knowledge in ethics is currently generated by scholarly introspection and deduction from the premises so derived. Knowledge is validated via review and debate within the academic community rather than through any carefully controlled confrontation between theoretical claims and empirical observations. Lofty statements about what makes for acceptable introspection are the norm. Study in such "underdeveloped" areas calls for careful epistemological judgements. Finding a place for empirical work in such fields can be daunting. To start, one must develop an argument about what constitutes relevant data against which to measure a claim of truth. The argument must also justify the epistemology used in making any truth claims.

Partially in response to this need, experimentalism is slowly beginning to creep into the methodology of ethical discourse. Many of the new initiatives are being undertaken by scholars from outside the field. Beck's paper [1994], framed as an experimental study in behavioral economics, has important implications for the question of distributive justice. It is one more interesting paper to add to the small but growing bibliography of empirical work in this area. As in other areas of inquiry, experimental studies in ethics and political theory are driven by the theoretical frontiers of the field. In issues of distributive justice these boundaries are drawn by two bodies of literature: the Rawlsian (and other impartial reasoning theories regarding the nature of distributive justice) and the economic/behavioral science literatures. The impartial reasoning literature has tied distributive justice to a standard of fairness in income distribution based on a consensus of individual judgements arrived at under certain idealized conditions. The behavioral literature has emphasized

that social choices with regard to income distribution can be seen as choices among social insurance programs. Such choices focus on the behavioral response to risk, loss, and uncertainty. However, each body of literature contains elements of the other's arguments.²

Given the underdeveloped role of controlled observation in ethics, claims to knowledge tend to be evaluated within a classical "coherence" framework.³ That is, early findings are assessed with regard to how well they seem to fit with what is accepted as knowledge. And there is little firm ground which can be used to assess the import or content of any particular *new* "finding" which violates conventional wisdom, even if it is well tested in a series of careful experiments. This is because the theoretical concepts are not usually framed in such a way that a particular empirical operationalization can be argued to "capture" them accurately and so one cannot be sure that the data from the experiments is directly relevant to the theoretical claims being made. Hence, until more exacting standards for the application of data to theory are developed, the main repository for accepted theory in such areas can be expected to be the underlying consensus (built on common sense and introspection). New findings resulting from different methodologies may be properly viewed as empirical probes of the theoretical base of the discipline. One consequence of this is that explanatory and theoretical formulations based on empirical findings which fly in the face of such "common sense" will be open to critical questioning. This tendency is reinforced by the fact that different experimental designs, testing a narrow set of theoretical principles, are likely to furnish results that will conflict to some extent (although they may be mutually supportive in many others). Proponents of the existing order are to be expected to seize upon differences in findings to uphold the prevailing orthodoxy.

This is the context in which Beck's [1994] experimental study of individuals' preferences for the distribution of income offers some interesting insights. It casts light on both substantive questions "What do individuals prefer regarding income distributions?" and "How are these preferences related to justice in distribution?" and on the two methodological questions "How can we come to know those preferences?" and "How can observations of preferences be used to justify a normative claim?" Given our own extended experimental program in this area [Frohlich and Oppenheimer, 1992], his work provides both confirmatory evidence for our main findings, and conflicting evidence regarding some of our secondary results. Reflecting on the similarities and differences in the two designs and findings provides insight into the problems of using experimental methods in this area of inquiry, and, indeed, in experimental economics in general.

COMPARATIVE FINDINGS

Beck's two main explicit conclusions are that individuals' preferences for an income distribution can be explained by individual risk propensities and that subjects reject Rawls's difference principle (presumably because, although they are risk averse, they do not exhibit the extreme risk aversion Rawls posits). Their risk

aversion leads individuals to choose a distribution which offers them an income support level: some floor below which their income will not fall. Another of his conclusions is a methodological message for ethicists. Implicit in his paper, and central to our own work, is the conclusion that one can both create situations in a laboratory which induce individuals to reason impartially about income distributions and obtain consistent results in those contexts.

Our own findings, that individuals will time and time again, in many settings and cultures, choose to set a minimum income floor for themselves and for groups of which they will be a part, support his finding that individuals choose a floor, or minimum income. The average value of the floor incomes chosen in our 52 North American experiments was \$10,860. When Beck's results are analyzed with an eye towards the average minimum level (in the binary income distributions) chosen by his subjects, a surprisingly comparable result emerges. The agreed upon average tax rates in his three experimental treatments are 0.2964, 0.2909, and 0.2917 for an overall average of 0.2930. Given his calculation formula, this implies a minimum "income" from the experiments of \$10,713!

Our results do not, however, support his argument that individual preferences for and group choices of income distribution are mainly a function of attitudes regarding risk.

In both Beck's work and ours, subjects chose from among risky alternative income distributions. But Beck structured a much simpler set of choices and framed the decision contexts in a number of different ways. His design asked individuals to choose from sets of risky binary income distributions, first as individuals and then for, and in, groups of varying sizes. Each of his experimental groups faced only two possible payoffs, and in only one experimental group was the choice framed explicitly in terms of an average annual income. We, on the other hand, placed groups of five experimental subjects in situations in which they had to choose from among principles for income distribution for hypothetical societies that they would subsequently experience. Our design induced a degree of impartial reasoning within the subjects. In our experiments subjects could not know, while deliberating, to which income class they would later be randomly assigned. Thus they were forced to take into account the welfare of all income classes equally.

Under these conditions of impartial reasoning, a consensus appeared relatively easy to reach. In all five locations, encompassing four cultures,⁴ where our experiments were run, consistent results were obtained. The vast majority of groups chose a principle which maximized the average income subject to a constraint that everyone would be guaranteed a certain minimum (floor) income. Yet the desire to impose a guaranteed floor seems to have come into play in his experiments in a way similar to the way it entered ours. Our findings are consistent with his which, on average, identify a floor, or minimum income. On the other hand, in both experiments, Rawls's principle is almost universally rejected and (in ours) was usually ranked as the least desirable.

Considering how similar our resulting floors are, it should be mentioned that these floors correspond to what most would view as one function of the modern welfare state: to deliver a minimally viable level of income support. But that interpretation may be stretching Beck's results, because the figures we are referring to are his aggregate results. In his study, more so than in ours, there is considerable variance in the choices of the individual groups. Further, he has three separate treatments to which 6 groups were subjected. Only one of those groups was presented with choices explicitly framed as annualized incomes. In the rest, the choices were framed in terms of individual dollar payoffs. This contrasts with our experiments, in which choices were in terms of annualized incomes. Not only were our samples broader (from several locations and countries) but our results had considerably less variance. Moreover, under supplementary analysis (to be discussed below) our results seem to point to a coherence with a level of support which reflects societal norms.

A major divergence between the conclusions of our studies is Beck's assertion that only individual risk aversion seems to be related statistically to the choice of the height of the income floor for the group. His operationalization of risk aversion was built directly into the binary income distributions from which subjects were asked to choose. The choice of a particular binary distribution "revealed" the individual's risk preference. It might be argued that Beck's first two treatments — one involving an individual choice for oneself, the other an individual choice for the group — both capture only risk preference. An individual in the latter case might well be interested in the welfare of others, but in projecting him or herself into the others' shoes in trying to decide what risks would be preferable, there is no means for any substantively new empirical data to enter into the individual's decision. Beck's design does not provide the means for informed impartial reasoning regarding these concerns. Thus, it is reasonable to assume that the individual would project his or her own risk preference onto others, and would not be able to integrate, into the reasoning, meaningful information about others' preferences. Thus, one would expect to find the same result in both cases: both being the reflection of one individual's preferences over risk. One might also conclude that the mean results of the third treatment should be the average of the individual preferences in the first two treatments. Indeed, that is what was found. In other words, Beck's experiments might be viewed as operationalizing only risk-preferences.⁵

We, on the other hand, structure our experiments around an attempt to invoke informed impartial reasoning on the part of subjects precisely because of an interest in the normative aspects of income distribution. The reason normative inquiry fits with decisions arrived at from an impartial point of view can be easily explained. Impartial decisions have a claim to ethical validity. The invocation of impartial reasoning is often argued to be a precondition for ethical introspection and moral knowledge. As far back as the 1st Century B.C., Publius Syrus of Rome noted that when disputes arise, there is an inevitable problem of bias. His dictum to avoid this problem and get a fair settlement was, "No one should be judge in his own case." For a judgement to be fair the interests of all must be taken into account evenhandedly.

One must consider the situation impartially. Many suggestions have been made regarding how this might be done in principle. Henberg notes some ways in which this impartiality might be accomplished "... C. I. Lewis, for instance, suggests that impartial valuations are best rendered by imagining that the experience of all concerned persons were one's own ... '... as, for example, if you had to live the lives of each of them seriatim.' Similarly, Richard Hare suggests that people who render differing moral judgements should imagine that their desires and inclinations are exchanged for the desires and inclinations of their antagonist" [1963, 723].⁶

The intuition behind these, and many other similar positions is simple: To be fair, we are to project ourselves into the position of the relevant other(s) and give equal weight to their concerns. The judgement about what is fair is to take place from an impartial point of view.

But how can impartiality be achieved in practice, as opposed to, in principle? A clue is given by the well-known device for dividing a piece of cake fairly between two individuals. One person divides the piece and the other chooses. The cutter is thereby given *every incentive to be as fair as possible in the division* because (s)he has to take into account the interests of both parties in the division. (S)he is rendered effectively impartial. So "you cut, I choose," is a universally acceptable procedural solution to the two-person problem of fair division of a fixed quantity.

John Rawls [1971] proposed generalizing this procedure to choices of rules for governing society. He envisioned a kind of thought experiment. We are to imagine individuals (representative of classes and other social positions) deprived of all *particular* information about their tastes, talents, dispositions, etc. and asked to play the roles of judges regarding rules to govern a society, which they and their progeny are to inhabit. The lack of particular knowledge prevents them from knowing what role they are to play in that society. This imperfect information induces conditions that generate impartial reasoning.

By requiring subjects to choose a set of payoffs with the knowledge that they (and possibly others) will subsequently be randomly assigned to one of those payoffs, both Beck and we create conditions under which individuals have incentives to reason impartially. The main explicit findings of the Beck paper (that individuals' preferences over income distribution are primarily a function of individual risk aversion) were built on explicit choices regarding risky alternatives of income flows. Our subjects chose income distribution from an impartial point of view only after a group discussion was held and all members had an opportunity to express their preferences and to take into account others' expressed preferences.

In our operationalization one measure of the risk aversion of our subjects was derivable from the risks implicit in our distributions⁷, but a perhaps less sophisticated, more direct operationalization was also contained in a questionnaire administered to all subjects. Using the latter independent measure of individual (rather than group) preferences we found that risk aversion was related to neither the preferences and choices of the individual subjects nor the outcomes chosen by the groups. Further, we found that framing the situation in terms of gains as opposed to losses, or as normative, rather than self-interested, insurance decisions, had no statistically

significant effect. Since such framings have been found elsewhere [Tversky and Kahneman, 1981] to be related to affect choices as a function of risk preferences, we explored what alternative underlying factors might have accounted for the variance in floor choices of different groups. We found that a cluster of individually held political values regarding equality, redistribution, and the legitimate role of government in these matters explained much of this variance (43 percent in 61 experiments, [Frohlich and Oppenheimer, 1992, 89]).

Of course given the considerable differences between the two studies, it is difficult to assess either what generated, or what to make of, the similar findings regarding the height of the floors, and the potentially differing role of risk aversion. In our experiments the content of the discussions, the debriefing questionnaire, and the statistical analysis indicate that the choices stem from an ethical concern, and not from simple risk aversion. An analysis of the factors underlying our subjects' preferences revealed that the balancing of three conflicting ethical concerns led people to choose an income floor as a principle of just distribution. These three normative principles are: (1) *Need*: no one should be allowed to starve or to live in absolute misery, (2) *Just deserts or Entitlement*: people should be rewarded for effort and productivity and (3) *Efficiency*: the income distribution should not contain incentives which reduce productive activity. An income floor assures (1) and the relatively low level of the floor tends to limit threats to (2) and (3). But we can explore further some of the fundamental differences in how the two studies structured the research problem.

The first difference might be thought of as the problem of framing. The discussions in and orientation of our experiments dealt with the social *distributional* characteristics of income for the group. This is somewhat different from the concerns operationalized in Beck's effort. Without this focus on the pattern of distribution for a society or group, it is perhaps less surprising that individual risk preference came to be viewed as the explanation for the choices. Further, the characteristics of the distributions were quite different. In Beck's experiments there were two income classes, and each subject had an equal probability of being in either. Since the probability of class assignment was made with replacement, everyone had a chance of getting the higher income, (no one was *required* to be in the lower class). This may have led to different evaluative imperatives than in the structure which we employed. In our experiments subjects knew that there were five classes (when there was no production), and that each class would be filled. The lottery was without replacement.

Other differences in the experimental designs had to do with the imposition of a deadline in the decision making in Beck's experiments. We not only did not have an experimentally imposed deadline, there was an elaborate procedure required for ending the debate and recording the votes.⁸ This was structured to approximate Rawls's condition of a decision taken at a point of reflective equilibrium.

What then are we to say about these differences? What is the interpretation and importance of conflicting or corroborating results when experimental designs, and even motivational theories, differ? A first response is to invoke an argument based on novelty. The theories are young, and largely untested. At least one (Rawls's) was

developed without any intention or expectation that it ever be tested. Moreover, beyond Rawls and risk preferences there are a variety of other theoretical approaches which predict similar results regarding the choice of principles of distributive justice. So for example, utilitarian and rule utilitarian principles might well have similar implications to those of impartial reasoning and the risk-behavioral approaches. And since some of these differing approaches generate similar predictions, the fact that the findings of a variety of experiments and operationalizations substantiate similar generalizations may be of importance. Further, as the conflicts among the data sets are sorted out, their implications for the different theories may become clearer. So, for example, it would be interesting to discover whether Beck's conjecture that risk aversion plays a large role in the group's choice holds up when tested directly in other designs.

If we consider the big picture, it might be that Beck successfully framed the problem as one of individual risk taking. Individuals had no other consideration to bring into the picture. His subjects (assuming that his conjecture is correct) decided on the basis of preferences regarding risk. And perhaps we framed the issue as one of impartial reasoning regarding a social decision on income distributions. As such, we are back to the two theoretical paradigms mentioned at the beginning of this note. It is not at all clear what role an assumption of risk neutrality does or does not play in an impartial reasoning framework. Examining a variety of choices under conditions approximating impartial reasoning may be a means of teasing out a structure of risk preferences that one might want to label as "neutral" in a deep normative sense. It may be a means of identifying what representative ideal individuals would choose under truly ideal conditions of impartiality.

SOME IMPLICATIONS FOR EXPERIMENTATION

As noted in Frohlich and Oppenheimer [1992], there are many impediments to identifying *a priori* what imaginary individuals placed under hypothesized conditions of impartiality would choose. Many of the relevant details of the information and values that they bring to bear cannot be adequately specified and a determinate result cannot be generated. Our response to that indeterminacy has been to simulate (albeit imperfectly) the conditions which induce impartial reasoning in a situation involving choices of income distributions for an experimental society. By using real subjects, our objective was to attempt to get an estimation of what might be chosen as fair under ideal conditions of impartiality.

But pursuing that end has served to identify a tension in experimental economics which is less likely to show up in such sciences as experimental physics. We might call this tension the problem of representativeness.⁹ In disciplines such as physics, being able to explain how objects behave under laboratory conditions is a goal of direct interest. Concern with performance under non-laboratory conditions is typically viewed as an engineer's concern. Field testing of new devices and applications are usually considered to be outside the domain of most experimental branches of inquiry in the theoretical sciences.

This is obviously less the case in the social sciences. In the social sciences there is a desire to be able to discuss theoretical performance under field conditions. Laboratory experiments which employ very "clean" conditions, allowing for variation in only one variable across treatments, are often so pristine as to tell us nothing other than that individuals do indeed respond to incentives when there are no other concerns.¹⁰

The relationship of the results to the field conditions (the so-called external validity) is questionable. If the experiments are to be vehicles for the introduction of real world concerns, some richness of environment must be thoughtfully woven into the laboratory conditions. But the introduction of fuller texture into the experimental design poses other threats. Such a threat is reflected in Beck's observation that in his group decisions, a single individual could (and in one instance did) sway the group decisions. We noted the same finding. But what does this lead one to conclude? Certainly, in the real world individuals can and do take over leadership roles. What do we aspire to do in our experiments: to make a design which leads to an artificially egalitarian discussion or one which is more permissive and can lead to a discussion in which one individual dominates? This is a tension which must be thought about and addressed explicitly. It is possible to design experiments to observe such behavior more carefully so that we can induce some of the patterns we will have to include in our theorizing.¹¹

Of course, more complex experiments are harder to interpret, but they also permit one to induce more regarding the non-laboratory environment. For example, the current research agenda regarding dictatorship and ultimatum experiments implicitly takes a strong and clear position on the question of whether the experiments should be "clean" or "realistic." The choice is being made to opt for one extreme, toward creating environments which are pure tests of a particular and narrowly defined theory: self-interested microeconomic behavior. The risk of such a stance is that the results may be totally without external (to the laboratory) referent. This dilemma poses a question which goes to the heart of the experimental enterprise: in the social sciences are experiments to be solely a handmaiden to pure theory, or might they play a broader role? The issue of external validity is a theoretical and methodological problem which must be given some thought in the near future in experimental economics.

NOTES

1. It must be noted that a substantial body of opinion in the field, tracing its pedigree back to Hume's famous separation of fact and value, holds that empirical data *cannot* play any role in justifying ethical statements.
2. Howe and Roemer [1981] identify a theoretical link between the two approaches in an interesting way.
3. Such an epistemological stance is well developed by a number of authors, including Quine [1960, 1961] and Brink [1989]. This situation is characteristic of many fields of inquiry in their early stages.
4. We conducted experiments in Manitoba, Maryland, and Florida, and provided experimental protocols and consultations to experiments run in Poland and Australia [Frohlich and Oppenheimer, 1992; Lissowski, 1991; Jackson, forthcoming].

5. The exception might be the single treatment and group that discussed a group choice as framed in terms of societal incomes. But since only two income classes were involved, realistic concerns regarding real-world income distributions would be diluted by his treatment.
6. The internal citations are to C. I. Lewis [1946, 547] and R. M. Hare [1963, 123].
7. Beck notes, and we agree, that it is very difficult, if not impossible, to identify detailed risk preferences from choices of complex income distributions.
8. However, in some of the experiments done in Poland there was a deadline (of 5 minutes) and it was only under these conditions that any group was unable to reach a consensus. We also imposed no penalty for groups that could not reach a decision. Finally, all our groups were the same size, while Beck's groups varied in size (but note he found no statistically significant relationship between group size and the decisions reached).
9. Below we discuss the question of the fit between the laboratory and the field environments. There is another type of representativeness that needs consideration. Since we are interested in generalizations regarding all individuals the issue of the representativeness of the samples in the experiment is also crucial. And in issues of impartial reasoning (where all individuals are to consider their fellows) this type of representativeness takes on even greater importance.
10. Of course there are always an infinite number of variables that vary between experimental treatments. Most are not germane to the measurements at issue, and those that are, are usually dealt with by randomization of subjects to the different treatments.
11. But richly textured experiments may not be "clean." And, given current gate-keeping that may mean that they may not be publishable.

REFERENCES

- Beck, J. H. An Experimental Test of Preferences for the Distribution of Income and Individual Risk Aversion. *Eastern Economic Journal*, Spring 1994, 131-46.
- Brink, D. O. *Moral Realism And The Foundations of Ethics*. Cambridge: Cambridge University Press, 1989.
- Frohlich, N. and Oppenheimer, J. A. *Choosing Justice: An Experimental Approach to Ethical Theory*. Berkeley, California University Press, 1992.
- Gauthier, D and Sugden, R., ed. *Rationality, Justice and the Social Contract*. Ann Arbor: University of Michigan Press, 1993.
- Hare, R. M. *Freedom and Reason*. Oxford: Oxford University Press, 1963.
- Henberg, M. C. Impartiality. *Canadian Journal of Philosophy*, December 1978, 715-24.
- Howe, R. E. and Roemer, J. E. Rawlsian Justice as the Core of a Game. *American Economic Review*, No. 5, 1981, 880-95.
- Jackson, M. A Fair Share. *Journal of Theoretical Politics*, forthcoming.
- Lewis, C. I. *An Analysis of Knowledge and Valuation*. LaSalle, IL: Open Court, 1946.
- Lissowski, G., Tyszka, T. and Okrasa, W. Principles of Distributive Justice: Experiments in Poland and America. *Journal of Conflict Resolution*, March 1991, 98-119.
- Quine, W. V. O. *Word and Object*. Cambridge, Mass: MIT Press, 1960.
- _____. *From a Logical Point of View*. New York: Harper & Row, 1961.
- Rawls, J. *A Theory of Justice*. Cambridge: Harvard University Press, 1971.
- Tversky, A. and Kahneman, D. The Framing of Decisions and the Psychology of Choice. *Science* January 1981, 453-8.