Journal of Public Economics 92 (2008) 1811–1820

Contents lists available at ScienceDirect

Journal of Public Economics

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

Social preferences and public economics: Mechanism design when social preferences depend on incentives

Samuel Bowles [a,b,*], Sung-Ha Hwang [c,1]

[a] Santa Fe Institute, United States
[b] Dipartimento di Economia Politica, University of Siena, Italy
[c] Departments of Economics and Mathematics, University of Massachusetts at Amherst, United States

ARTICLE INFO

Article history:
Received 27 April 2007
Received in revised form 27 February 2008
Accepted 6 March 2008
Available online 21 March 2008

JEL classification:
D52 (incomplete markets)
D64 (altruism)
H21 (efficiency, optimal taxation) H41 (public goods)

Keywords:
Social preferences
Implementation theory
Incentive contracts
Incomplete contracts
Framing
Motivational crowding out
Ethical norms
Constitutions

ABSTRACT

Social preferences such as altruism, reciprocity, intrinsic motivation and a desire to uphold ethical norms are essential to good government, often facilitating socially desirable allocations that would be unattainable by incentives that appeal solely to self-interest. But experimental and other evidence indicates that conventional economic incentives and social preferences may be either complements or substitutes, explicit incentives crowding in or crowding out social preferences. We investigate the design of optimal incentives to contribute to a public good under these effects would make either more or less use of explicit incentives, by comparison to a naive planner who assumes they are absent.

1. Introduction

In his *Essays*: *Moral, Political and Literary* (1742) David Hume (1964):117–118 recommended that

> in contriving any system of government… every man ought to be supposed to be a *knave* and to have no other end, in all his actions, than private interest. By this interest we must govern him, and, by means of it, make him, notwithstanding his insatiable avarice and ambition, cooperate to public good.

Hume's maxim that public policies should harness self-regarding preferences to public ends remains a foundation of public economics. Its wisdom is buttressed by ample evidence that conventional incentive-based contracts and policies often work very well (Laffont and Matoussi, 1995; Lazear, 2000).

* Corresponding author. Santa Fe Institute, United States.
E-mail address: bowles@santafe.edu (S. Bowles).

[1] We would like to thank Margaret Alexander, Lopamudra Banerjee, Ernst Fehr, Duncan Foley, John Geanakoplos, Suresh Naidu, Seung-Yun Oh, Sandra Polania Reyes, Luc Rey-Bellet, Carlos Rodriguez-Sickert, John Roemer, Bob Rowthorn, Paul Seabright, Rajiv Sethi, Joaquim Silvestre, Peter Skott, Joel Sobel, E. Somanathan, Tim Taylor, Elisabeth Wood, Giulio Zanella and members of the Yale Law School Legal Theory Seminar for their contributions to this research. Thanks also to the Behavioral Sciences Program of the Santa Fe Institute, the U.S. National Science Foundation, the European Science Foundation and the University of Siena for financial support of this project.

0047-2727/$ – see front matter
doi:10.1016/j.jpubeco.2008.03.006

But Hume only "supposed" citizens to be knaves. In recent years experimental evidence has endorsed Hume's caveat (immediately following the above passage) that the supposition is "false in fact": altruism, reciprocity, and what the classicals called civic virtues are powerful and common motivations (Camerer, 2003; Fehr et al., 2007; Gintis et al., 2005). The empirical importance of other-regarding motives for public economics has also long been recognized and has recently been affirmed in studies of tax compliance (Andreoni et al., 1998; Pommerehne and Weck-Hannemann, 1996), political opinion and voting concerning income security and redistribution measures (Fong et al., 2005), generalized obedience to law (Kahan, 1997), and other areas critical to public economics.

Hume, Jeremy Bentham and the other classicals advocating self-interest as a basis of public policy design did not ignore the social preferences that underlie moral behavior. What Adam Smith termed "the moral sentiments" played a central role in their thinking. But they assumed that ethical motivations would be unaffected by incentive-based policies designed to recruit self-interest to public ends. Along with civic virtues, explicit incentives and constraints could thus contribute additively to good government. According to this view, taxes or subsidies affect individual utility and hence behavior only indirectly, that is by altering the economic costs and benefits of the targeted activities. These and other explicit incentives thus do not appear directly in the citizen's utility function. As a result, the behavioral effects of moral sentiments and the material interests are separable, the effects of each being independent of the levels of the other. But when separability does not hold, the two kinds of motivations may be either complements – social preferences being heightened by incentives appealing to self-interest – or substitutes, when explicit incentives are said to crowd out social preferences.

A consequence of the classicals' implicit 'separability assumption' is that they failed to take account of how harnessing self-interest to the public good might either compromise or enhance civic virtues. While in contemporary economic theory separability is not explicitly assumed and could be abandoned, modern public economics, mechanism design and related fields continue the classicals' practice. However a great many experiments and observations in natural settings suggest that social preferences are often important influences on behavior, and that the salience of these preferences varies with the kinds of explicit incentives that are implemented.

If the separability assumption is false, policies designed on its basis will generally be non-optimal, and explicit incentives will be over-used or under-used. Over-use of explicit incentives when crowding out is the case was the central theme of the study of blood donations by Richard Titmuss (1971). In similar vein Albert Hirschman (1985): 10 castigated economists who propose "to deal with unethical or anti-social behavior [solely] by raising the cost of that behavior… [because they] think of citizens as consumers with unchanging or arbitrarily changing tastes" adding that "A principal purpose of publicly proclaimed laws and regulations is to stigmatize anti-social behavior and thereby to influence citizens' values and behavioral codes." The implications for constitutional design of cases in which "institutions themselves affect preferences" were first developed by Michael Taylor (1987): 177 and subsequently expanded by Bowles (1989), Frohlich and Oppenheimer (1995), Kreps (1997), Frey (1997), Bowles (1998), Cooter (1998), Ostrom (2000), and Bar-Gill and Fershtman (2005).

The economic intuition underlying Titmuss' and Hirschman's concerns is that because crowding out reduces the effectiveness of explicit incentives, they would be used less by a sophisticated social planner cognizant of the crowding out problem, by comparison to a naive planner, namely, one who assumes that economic and moral motives are separable. If crowding out is so strong that the incentive has an effect the opposite of its intent, this is of course the case. But the effect of crowding out need not be literally counterproductive in this sense; and where the effectiveness of incentives is blunted but not reversed, the implications for the optimal use of incentives are far from obvious. The reduced effectiveness of the incentive associated with crowding out would entail a *larger* incentive for a planner designing a subsidy to ensure compliance with a quantitative target, a given fraction of the population receiving anti-flu injections for example. We will show that these seemingly conflicting intuitions are both correct. To do this we develop a model of optimal explicit incentives in the presence of both crowding in and crowding out, and use the model to identify cases in which crowding out entails greater or lesser use of incentives by the sophisticated planner.

To analyze these cases we will ask what incentives would be adopted by a social planner who wishes to maximize the aggregate utility of citizens (by "incentives" without adjective we mean those appealing to conventional self-regarding preferences). We will say that incentives are over-used if the sophisticated planner who takes account on non-separability would adopt a lesser level of incentive than would the naive planner, and conversely.

In the next section we survey the empirical literature on non-separability. We then introduce a model of public incentives when individuals with social preferences may contribute to a public good, using this model to clarify the separability assumption and how it may be violated. In Section 4 we use the model to show that the sophisticated social planner seeking to ensure a target compliance level of contributions by citizens will implement a higher level of incentives (or none at all) if crowding out holds. In Section 5 we study optimal incentives for the sophisticated planner who maximizes total social welfare, taking account of the values of the citizens both as components of their utility and influences on their behavior. We find that in a public goods setting, as in the compliance case, the sophisticated planner may make more use of incentives than the naive planner when crowding out is the case. The economic intuition behind this surprising result is evident in the compliance case. In cases where the marginal social benefits of the public good rise sharply as the shortfall from its socially optimal level increases (a limiting case of which is the target compliance problem), the fact that crowding out makes the incentive less effective requires its greater use. Where decreasing marginal returns to the public good are modest or absent the sophisticated planner will make lesser use of incentives when crowding out holds, as expected. In Section 6 we consider the implications of non-separability for public economics.

2. When separability fails; evidence and explanations

The underlying social and psychological mechanisms accounting for non-separability include the following (Bowles (in press) and Frey and Jegen (2001) survey the experimental and other evidence).