

Cambridge University Press

978-0-521-33045-9 - Information, Incentives and the Economics of Control

G.C. Archibald

Excerpt

[More information](#)

Part I

Introductory

Cambridge University Press

978-0-521-33045-9 - Information, Incentives and the Economics of Control

G.C. Archibald

Excerpt

[More information](#)

1

Two preliminary matters

1.1 Individualism and holism

It has long been usual to find, at the beginning of a work on welfare economics, a statement to the effect “I adopt the liberal principle that individual preferences are to count: take it or leave it.” I indeed adopt this principle here; but I think that some justification may be in order. Nothing like a complete justification can be attempted: that would require a major work of political philosophy. Nonetheless, we may consider one alternative to individualism, and some difficulties.

In common speech, we often use collectives such as “France,” “the working class,” or the “elderly.” These collectives may be – perhaps usually are – employed simply as shorthand for aggregates of individuals. They may, however, mean more than this. It may be believed that the collective, the group or “whole,” is an entity, and actually exists in its own right. Philosophers call this view “methodological holism.” There are many versions of holism. The version most obviously antithetical to individualism was identified by Popper (1957), and most sharply defined by Agassi (1960). The key is his Proposition 4: “If ‘wholes’ exist, then they have distinct aims and interests of their own.” This is, perhaps, frightening. A holist in this sense may talk, for instance, of the “interest of the state,” or the “national interest,” without at all intending by these terms merely a shorthand for certain collections of individuals. The liberal alternative adopted by most economists is, of course, “methodological individualism.”

Cambridge University Press

978-0-521-33045-9 - Information, Incentives and the Economics of Control

G.C. Archibald

Excerpt

[More information](#)

4 **Introductory**

It is not my purpose to try to persuade on this matter. I shall say only that if the choice is between individualism and holism as defined by Agassi, perhaps few of us will hesitate. We should, however, be aware that we are making a *choice*, and that it is perhaps as difficult to give a rational argument for such a philosophical choice as for any faith. We should also notice that we have to be very careful about the meaning of some convenient terms, such as “the interests of the Third World.” We need not, however, become so timid as to eschew entirely the use of collective terms: whether the intent is holist or individualist is usually evident from the context. (I say “usually,” not “always.” Whether or not Marx, for instance, was a holist in Agassi’s sense may not always be entirely obvious.)

If one adopts individualism as one’s methodology, choices still remain. The social scientist may adopt it for both positive and normative purposes. If he adopts it for both, he immediately faces another problem: is it *sufficient*? Is he, that is, willing to become a monist in his ethics, believing not only that individual preferences count, but that *only* they count, or should count, in decisions on economic or social policy? Clearly, the choice of individualism over holism does not itself entail monism. Nonetheless, one monist philosophy commands attention. The methodological individualist has somehow to deal with the problem of aggregating individual preferences. Utilitarianism offers a comprehensive and sufficient solution to that problem. The individualist who is not a utilitarian really has no solution to offer, as we know from the work of Arrow (see particularly his 1951b). Yet there are individualists, of whom I am one, who cannot accept monist utilitarianism, and must be content with what Brian Barry has somewhere called a “pluralist cocktail” of ethical principles (which does not entail giving no weight to any utilitarian argument).¹

This is, of course, not a book on ethics, and much of the discourse is strictly positive. Yet in even a work on the implementation of welfare economics rather than on welfare economics *per se*, it seems that some reason for taking the subject seriously, or at least advertising the author’s methodological choices, may be appropriate. And when we consider extended

preferences, in ch. 2 below, we shall find that we may want to ask if some preferences should *not* count.

1.2 Incentive compatibility

It is consistent with a positive individualist methodology to assume that economic agents act entirely, or mainly, for motives of self-interest, although this is not entailed. Indeed, in ch. 2 I investigate extended preferences – assuming, that is, that the agent explicitly takes into consideration the wellbeing of some, or all, of his fellow citizens. We shall find that, on apparently quite “reasonable” restrictions, these preferences are perfectly consistent with standard “liberal” results and policies: agents will for the most part behave as ordinary selfish maximizers. This in turn implies that, in considering any methods for “control” of the economy, incentive compatibility must be taken seriously.

We owe to Adam Smith the insight that matters go more smoothly if institutions are such that private and social interest coincide. D.H. Robertson (1956) put it clearly. “What do economists economize on?,” he asked. This was not a rhetorical question. His answer was: Love. He explained that love is scarce, and that it is wasteful to depend on it for everyday social arrangements that depend, or can be made to depend, simply on self-interest. As Smith (1776) put it “It is not from the benevolence of the butcher, the brewer, or the baker that we expect our dinner, but from their regard to their own interest. We address ourselves not to their humanity but to their self love, and we talk to them not of our necessity, but of their advantages” (p. 14).

If we economize on love, we do more: we economize on policemen. If it is in the interest of agents to do what is socially desirable, we have neither to appeal to their altruism nor employ policemen to ensure their good behavior. Institutions that economize on policemen also economize on something else expensive: information. If it is in agents’ interests to “do the right thing,” there is no need to use resources to find out just what they are doing, or how.

6 **Introductory**

The standard formulation of the principal–agent problem is precisely as a problem in economizing on information, love, and policemen. It is assumed that all concerned are exclusively self-interested, and that lack of information entirely precludes monitoring of the agent. Yet there is a problem here. May not an agent himself become the principal in some subsidiary contract(s) that tend to subvert the object of the original contract (see Eswaran and Kotwal, 1984)? The possibility of side-payments suggests that agents may indeed become principals, and vice versa: the old *quis custodiet* question leads to an infinite regress.² I do not presume to offer any general conclusion on this matter. Notice that *any* social institution, existing or proposed, has at least an implicit incentive structure which requires examination, usually more for its unintended than for its intended consequences. And suppose that we do encounter, if not a demonstrably infinite regress, at least a tediously long chain of possibilities for side-payments and subversion: what do we do? Sooner or later, exhaustion sets in. We may also notice that in any such chain, perhaps at the first step, we shall encounter conduct regarded in many societies as immoral, and possibly illegal. If our object is to economize on policemen, that is not a sufficient excuse for terminating our enquiry: if the incentives to “misbehavior” are large enough, the jails will not be; and, in any case, policemen are but agents, and agents who may become principals.

The policy I have followed in this book is to pursue the possibilities of strategic behavior, and of side payments (agents becoming principals) as far as my own ingenuity and energy permit (and obviously no further). In at least two places, I have had to give up, and appeal for criminal sanctions. I can only warn the reader to be alert to possibilities that I have overlooked, or inadequately investigated.

2

Extended preferences¹

2.1 The axiom of selfishness and the Two Theorems of Welfare Economics

The preferences attributed to individuals in welfare economics are usually assumed to satisfy the *axiom of selfishness* – that is, each individual is assumed to order consumption bundles for himself without regard to anyone else's preferences or actual consumption, and is said to be better off if he receives a more preferred bundle. There are two reasons for doubting if this is a satisfactory foundation for individualistic welfare economics. The first is empirical: it is doubtful if people are, always and everywhere, so purely selfish. The second is that it is hard to find much force in normative prescription for a world in which all agents are, by assumption, amoral. (The difficulty of a utilitarianism that accepts the axiom as descriptively accurate but goes on to recommend policy on utilitarian moral grounds is well known.) It therefore seems worth trying to relax this axiom if we can: we might gain in positive content *and* add moral force to normative individualistic prescription. We may indeed drop it, but we must enquire into the cost of doing so, and with what we may replace it.

Since Arrow (1951a), two outstanding contributions by Edgeworth (1881) have commonly been called the First and Second Theorems of Welfare Economics. The First Theorem is that any competitive equilibrium is a Pareto-optimum. The Second is that any optimal allocation can be supported by competitive prices if the initial endowment is appropriate.

8 **Introductory**

There is no occasion to delay now to set out all the assumptions of the Edgeworth–Arrow theorems, some of which will be extensively discussed below. The immediate question is the effect that dropping the axiom of selfishness may have on these two theorems. The First Theorem is an obvious and immediate casualty. If economic agents are concerned with each other's welfare, there clearly may be competitive equilibrium, corresponding to some initial endowments, which are not optimal: they are not regarded as morally unacceptable merely by the outside observer, but by the agents themselves. So, can we “save” the Second Theorem? This is a very serious matter. We rely on this theorem for the notion that matters of equity and of efficiency may be considered separately, or “divorced.” We rely on it for any notion that competition is efficient or “good” (at least in a strictly convex economy). Perhaps we rely on it overmuch, even in a convex economy, since it may be argued that deliberate or purposive redistribution cannot be lump-sum, that the idea is inherently self-contradictory (unless, perhaps, it can be based on observable but immutable individual characteristics). However this may be, the immediate question is what restrictions are required on preferences if they are to be extended (or interdependent) *and* it is still to be true that an optimum allocation can be supported by competitive prices (given, of course, the other necessary assumptions).

2.2 Edgeworth's treatment of extended preferences (1881)

This question has been investigated before – first, indeed, by Edgeworth (1881) himself! Edgeworth's work does not seem, however, to be very well known (it certainly has not reached the textbooks), and the subsequent history is rather diffuse. I therefore think it worthwhile to retell some of the story, using the opportunity to clarify some issues and correct some errors. Edgeworth considered the possibility of “sympathy” between economic agents. At one extreme, there is none: the neighbor's utility counts for nothing (the axiom of selfishness is satisfied). At

the other extreme, the neighbor's utility "counts for one": it is as important to the agent's happiness as his own (the Purely Universalistic case, as Edgeworth called it). In between, the neighbor's utility will "count for a fraction." Edgeworth, as a utilitarian, was able to represent this, in the two-person two-good case, with the additively separable utility functions

$$U^A(x_A, y_A, x_B, y_B) = \alpha u^A(x_A, y_A) + (1 - \alpha) u^B(x_B, y_B) \quad (2.1)$$

$$U^B(x_A, y_A, x_B, y_B) = \beta u^B(x_B, y_B) + (1 - \beta) u^A(x_A, y_A) \quad (2.2)$$

with

$$0 \leq \alpha, \beta \leq 1$$

and, if we wish to draw a box, the constraints

$$x_A + x_B = x$$

$$y_A + y_B = y.$$

(I use here neither Edgeworth's notation nor that of Collard, 1975, although the latter very opportunely reminded us of Edgeworth's contribution.²) Here U^A and U^B are the individual's "grand utility functions": their arguments are all the elements of the complete allocation of all goods to all members of society. Since, in this formulation, their arguments are the private utilities of all agents (functions of their own private bundles), we may call U^A and U^B the individuals' social welfare functions. It does not seem empirically unreasonable to suppose that individuals are moral agents and have opinions about social welfare which may be represented in some such way. Before asking if there are "better," or less purely utilitarian, ways of representing the social preferences of individuals, it will be convenient to examine some properties of Edgeworth's representation.

First, as Edgeworth stated, it does not disturb or distort the contract curve we should obtain if we assumed the individuals to be, in fact, selfish, and drew the contract curve for the case $\alpha = \beta = 1$. To see this, assume that neither agent is entirely selfish, setting $0 < \alpha, \beta < 1$. Now, maximizing either of (2.1) or (2.2), subject to the other reaching some preassigned value

Cambridge University Press

978-0-521-33045-9 - Information, Incentives and the Economics of Control

G.C. Archibald

Excerpt

[More information](#)10 **Introductory**

within the limits set by the quantity constraints, and rearranging the usual first-order conditions, we find the condition

$$u_1^A/u_2^A = u_1^B/u_2^B \quad (2.3)$$

(where the subscripts denote partial derivatives with respect to the arguments by order). Here α and β have dropped out. We have the same equal MRS condition that we normally obtain in the selfish case, where $\alpha = \beta = 1$. Collard (1975) (and see his 1978) calls this the “no-twisting theorem.” It is a remarkable result. Before going on to investigate the second of Edgeworth’s claims (the “contraction” of the contract curve), we explore this further.

We have the striking result that interdependent preferences, in Edgeworth’s representation, leave the contract curve – or at least part of it, whence the Second Theorem of Welfare Economics – undisturbed. How can this be?

Calculate, from (2.1), A ’s marginal rate of substitution between goods in B ’s bundle. It is

$$\frac{U_3^A}{U_4^A} = \frac{(1-\alpha)u_1^B}{(1-\alpha)u_2^B} = \frac{u_1^B}{u_2^B} = \frac{U_3^B}{U_4^B}, \quad (2.4)$$

that is, A ’s MRS between goods in B ’s bundle is B ’s MRS (and analogously, of course, for B ’s MRS between goods in A ’s bundle). That is why, given the equity considerations represented by $0 < \alpha, \beta \leq 1$, the Second Theorem holds. In the general, benevolent, case of $\alpha, \beta < 1$, each of U^A, U^B is increasing in u^A and u^B , which we may take to represent each individual’s “enjoyment” of his own consumption bundle. Yet the force of Edgeworth’s formulation is that each agent is concerned with the other’s wellbeing *only as that agent sees it*. Neither’s concern for the other’s welfare induces him to try to interfere with the other’s choice of (private) consumption goods. This is what Donaldson and I, in earlier work (Archibald and Donaldson, 1976a; 1976b; 1979) called the “non-paternalist condition.” It seems to represent J.S. Mill’s (1859) rule that one may not seek to coerce another individual for his own good, but only to avoid injury to a third party. “Coercion” here would mean attempted interference with the other agent’s choice of consumption bundle,