

Self-interest and sympathy in economic behaviour

Kjell Hausken

*Max-Planck-Institut für Gesellschaftsforschung, Cologne,
Germany*

Social science embodies a tension regarding the primacy of self-interest or sympathy as the actuating principle in human nature. There is considerable theory building in favour of or opposed to either principle. Economics, including game theory and principal-agent theory, has a traditional focus on the former. Some sociological, psychological and organization theories (Perrow, 1986) focus on the latter. Political science has a focus on both. These persisting states of affairs suggest a need to scrutinize the substantive roots of the controversy.

The theories of Hobbes (1651), Hume (1740; 1751), and Adam Smith (1759; 1776) are analysed here in game-theoretic terms. Presented in elementary terms, it is shown that game theory can account for both self-interest and sympathy. It is especially evaluated whether sympathy as an actuating motive in human nature gives rise to human interaction having characteristics other than interaction in models, like Hobbes', where such a motive has no influence. Analysing Hume's sympathy concept and Adam Smith's moral sentiments in such terms is intended to provide a richer soil for theory building with a possible joint focus on self-interest and sympathy, as well as enriching our general understanding of human interaction.

Sections two and three scrutinize those passages in Hume and Smith's works which are opposed to self-interest and accommodating to other-regarding behaviour displaying benevolence, sympathy and moral sentiments. The objective of these sections is to underline the negative aspects of self-interest and the positive aspects of sympathy. Section two shows how both Hume and Smith distance themselves from the notion of self-love as the sole actuating principle in human nature. Section three elaborates on Hume and Smith's interpretations of sympathy and moral sentiment as motives for action in general, and as motives for moral judgment. Having this solid foundation with emphasis on sympathy and a de-emphasis on self-interest, section four illustrates a model embodying both self-interest and sympathy, both egoism and altruism, contrasting Hobbes and Hume.

Section five is a transitional section, from focus on individual preferences in the preceding sections to focus on interaction between several players having these preferences in the sections that follow. Thus section five assesses the relative influence of self-interest and sympathy in Hume's political theory. Section six defines the convention concept in Hume's theory, illustrated with the battle of the sexes' game. Section seven illustrates three

games descriptive of political society: the assurance game, chicken game and the prisoner's dilemma.

Self-interest and sympathy

Section eight is an important section in the article, divided into three subsections. Having various kinds of individual preference and various political games in mind, the first subsection illustrates the static game men play in Hume's society. Special attention is given to how the presence of benevolence may alter game characteristics, possibly facilitating more cooperative behaviour. The second subsection incorporates the time aspect, crucial in Hume's theory. I analyse iterative games and construct one dynamic game describing interaction in Hume's society over time. The third subsection introduces more than two players to illustrate Hume's view of large societies as contrasted with small.

5

As we will see, sections up to and including section eight show the necessity of government, potentially to alter game characteristics so as to facilitate a more constructive society, and at best induce players to cooperative rather than exploitative or free-riding behaviour. Hence the penultimate section cursorily introduces the role of government, why it is necessary, and what influence it may have. The section ends with Hume's passages showing how men naturally attach moral approbation to artificially constructed systems in society.

Hume and Smith on self-love

Hutcheson, Hume and Smith are all critical of rationalism and egoistic theory; Hume especially is famous for his claim that "reason is, and ought only to be, the slave of the passions" (Hume, 1740, p. 415). Hume and Smith allow for a variety of motives for action, among them self-interest and sympathy. They are, however, both critical of one type of self-love which I analyse here in section two.

It seems that we can interpret Hobbes (1651) as not totally denying the existence of benevolence, pity and other principles in human nature not reducible to self-love, though in his political theory these principles seem to be absent. Hobbes' well-known basic notion is thus to assume "a perpetuall and restlesse desire of Power after power" (Hobbes, 1651, Ch. XI), defining power as "present means, to obtain some future apparant Good" (Hobbes, 1651, Ch. X). We can assume that Hobbes' political theory is solidly founded on the non-existence of motives such as benevolence and sympathy in human nature. Moreover, his theory seems to allow for at least a slight strain of malevolence which may flow from his account of man finding glory in seeking eminence over others, as we will see in section four.

Hume devotes more time to criticizing rationalism than egoism, presumably judging Hutcheson's critique of egoism as sufficient. But Hume has one short appendix on self-love. He says (Hume, 1751, p. 298): "The most obvious objection to the selfish hypothesis is, that, as it is contrary to common feeling and our most unprejudiced notions, there is required the highest stretch of philosophy to establish so extraordinary a paradox." Hume (1751, p. 297) criticizes Hobbes and Locke for "explain[ing] every affection to be self-love, twisted and moulded, by a particular turn of the

imagination, into a variety of appearances.” Hume’s alternative (1751, p. 297) is to “esteem the man whose self-love, by whatever means, is so directed as to give him a concern for others, and render him serviceable to society.” Hume gets a better grip on his critique when he goes on to argue that there are instincts in human nature, e.g. benevolence, which motivate us directly and are not reducible to self-love. He says (Hume, 1751, p. 301):

There are mental passions [and bodily wants] by which we are impelled immediately to seek particular objects, such as fame or power [or food], or vengeance without any regard to interest; and when these objects are attained a pleasing enjoyment ensues, as the consequence of our indulged affections. Nature must, by the internal frame and constitution of the mind, give an original propensity to fame, ere we can reap any pleasure from that acquisition, or pursue it from motives of self-love...Were there no appetite of any kind antecedent to self-love, that propensity could scarcely ever exert itself; because we should, in that case, have felt few and slender pains or pleasures, and have little misery or happiness to avoid or to pursue.

A Hobbesian would argue that Hume is not necessarily correct here since reputation or fame, for example, can be considered as means to one’s own happiness, welfare, etc. In fact Hobbes (1651, Ch. X) defines at least three kinds of reputation (for power, love, prudence, etc.) as power; and power is a means to more power. But how, then, has it come about that reputation has been considered as beneficial in the first place? Hobbes might here attempt to argue (Hume and Smith perhaps less probably) that we have just happened tacitly to agree on this, due to man’s needs, the scarcity of the world, etc. The way men interact when pursuing satisfaction of needs, something necessarily has to prove valuable as a means to one’s own happiness, like various kinds of reputation, whereas something else turns out to have less value. Using Smith’s theory (1759, p. 179) we can elaborate on this:

But that this fitness, this happy contrivance of any production of art, should often be more valued, than the very end for which it was intended; and that the exact adjustment of the means for attaining any conveniency or pleasure, should frequently be more regarded, than that very conveniency or pleasure, in the attainment of which their whole merit would seem to consist, has not, so far as I know, been yet taken notice of by any body.

Hence useful means are valued first for the ends at which they aim, but subsequently we are charmed by the beauty of their own sheer efficiency, thereby leading us to value means for their own sake. A Hobbesian may argue, alternatively to Hume’s claim (1751, p. 301) that “nature must, by the internal frame and constitution of the mind, give an original propensity to fame”, that if fame proves an efficient means to one’s own happiness, then propensity to fame is not necessarily implanted by nature, but is rather an *ex-post facto* consideration due to how men interact. We are in an equilibrium, an interlocked situation from which no one unilaterally and beneficially can deviate. From this it is still an open question whether self-interest and/or sympathy are originally implanted by nature or are secondary considerations.

Smith seems to judge Hume's critique of rationalism as mainly sufficient, and concentrates on criticizing egoism. We will here leave the so-called "Adam Smith problem" aside, and focus on how Smith argues against the possibility of deducing sympathy from self-love. He says (Smith, 1759, p. 31):

When I condole with you for the loss of your only son, in order to enter into your grief I do not consider what I, a person of such a character and profession should suffer, if I had a son, and if that son was unfortunately to die: but I consider what I should suffer if I was really you, and I not only change circumstances with you, but I change persons and characters. My grief, therefore, is entirely upon your account, and not in the least upon my own. It is not, therefore, in the least selfish. How can that be regarded as a selfish passion, which does not arise even from the imagination of any thing that has befallen, or that relates to myself, in my own proper person and character, but which is entirely occupied about what relates to you? A man may sympathize with a woman in childbed; though it is impossible that he should conceive himself as suffering her pains in his own proper person and character. That whole account of human nature, however, which deduces all sentiments and affections from self-love, which has made so much noise in the world...seems to have arisen from some confused misapprehension of the system of sympathy.

On the other hand, Smith writes, with regard to Mandeville, that the latter's system "could never have imposed upon so great a number of persons, nor have occasioned so general an alarm among those who are the friends of better principles, had it not in some respects bordered upon the truth" (Smith, 1759, p. 313).

Neither Hume nor Smith is very clear on the relative weighting of self-interest and sympathy as motives, and some of this controversy clearly is a dispute about words. However, there has been one discovery after Hume and Smith's lifetimes which seems to have received little attention, but which may throw light on our problem. Hume may be wrong in writing that "the nature of the subject furnishes the strongest presumption, that no better system will ever, for the future, be invented, in order to account for the origin of the benevolent from the selfish affections, and reduce all the various emotions of the human mind to a perfect simplicity" (Hume, 1751, p. 298). The discovery is that of Nietzsche (1886, aphorism 13) who assumed that the cardinal instinct of an organic being is "above all to discharge its strength".

We may conceive of human nature as having a high number of affects and passions, which are constantly in a war against each other, each trying to overcome the other, and each succeeding at each point in time, depending on the state of the organism – that is, its inclination, habits, past experience, sleep, nutrition, etc. Spinoza argued, not too foreign to the spirit of Hume and Smith, that "an affect cannot be restrained nor removed by an opposed and stronger affect...No affect can be restrained by the true knowledge of good and evil insofar as it is true, but only in so far as it is considered as an affect" (Hirschman, 1977, p. 23). Hume (1751, p. 299) argues: "The affections are not susceptible of any impression from the refinements of reason or imagination; and it is always found that a vigorous exertion of the latter

faculties, necessarily, from the narrow capacity of the human mind, destroys all activity in the former.”

Assume a set of affections, somewhat different in different persons, and different in a given person at different points in time. Nietzsche’s argument is that whatever drives and affects there might be in human nature, they are all of the same kind, in the sense that they all have a desire to discharge themselves and to overcome each other. Nietzsche’s well-known conclusion is to assume one single principle involving discharge as such – that of strength or the will to power. The nature of this process of continued discharge is such that will to power overcomes itself. If Nietzsche is correct, such instincts as self-interest and sympathy have it in common that they seek to discharge themselves, each according to its own peculiar nature. This may throw light on our problem, but it does not resolve whether one instinct is reducible to the other. It reduces instead both principles to one involving discharge.

It seems plausible to assume both self-interest and sympathy as actuating drives of human nature. Assuming sympathy to be completely reducible to self-interest implies arguing that for generations, mankind’s relationship to language, especially with regard to the use of such words as “sympathy”, has been imbued with a certain degree of falsehood. There may be superfluous words in our language, but we may argue that “sympathy” is likely to be too basic to be one of them. However, the weight and relevance of self-interest and sympathy may vary from person to person, and for each person from time to time. For example self-interest may prove explanatory in economic affairs, as also is *sympathy* when helping our neighbour’s young daughter to cross the street.

Hume and Smith on sympathy in moral judgement

Since both Hume and Smith attach such weight to sympathy, this third section analyses how they conceive of sympathy and moral sentiment. It is crucial to observe that motives for action in general and motives for moral judgement are kept apart. Unfortunately neither Hume nor Smith is very clear about this distinction. For action in general, sympathy, self-interest

Theme	Hume’s view	Smith’s view
Definition of sympathy	Sharing of pleasure or pain produced in a person affected by an action	Sharing of any feeling, especially focusing on the person’s motives, as a response to the particular situation of the other person
Moral approval of natural virtue	Spectator takes sympathetic pleasure in happiness produced by, e.g., benevolence	Spectator shares by sympathy the gratitude evoked by, e.g. benevolence
Moral approval of artificial virtue	Depends indirectly on utility, and ultimately on sympathy with the happiness of society	Regarding justice, depends on sympathy with resentment for harm
Social utility	Crucial regarding moral approval	Is an afterthought, not a foundation of moral approval

Table I.
Hume and Smith on
sympathy in moral
judgement

and other motives may prove actuating. Smith (1759, pp. 314-42), in reviewing earlier philosophical systems, elaborates on how the principle of approbation in various moral systems may be considered as deduced from self-love, reason or sentiment/feeling, the latter divided into moral sense and sympathy. Sympathy is crucial in Hume and Smith's explanations of moral judgement. Table I gives a sketch of these two authors' different conceptions of sympathy and moral sentiment.

Regarding sympathy in general, Hume (1740, p. 576) argues: "No passion of another discovers itself immediately to the mind. We are only sensible of its causes or effects. From these we infer the passion: and consequently these give rise to our sympathy." Hume (1740, p. 575) continues: "The minds of all men are similar in their feelings and operation, nor can any one be actuated by any affection, of which all others are not, in some degree, susceptible." Thus Hume, contrary to Hobbes who does not consider sympathy as actuating in human nature, seems to be of the opinion that our ability to sympathize is a natural inclination of our constitution, presumably biologically rather than culturally derived. Hume (1740, p. 618-19) continues: "...sympathy is a very powerful principle in human nature...sympathy is the chief source of moral distinctions...a sense of morals is a principle inherent in the soul, and one of the most powerful that enters into the composition". Hume argues that sympathy is a principle "which takes us so far out of ourselves, as to give us the same pleasure or uneasiness in the characters of others, as if they had a tendency to our own advantage or loss" (1740, p. 579).

Formalization of egoism and altruism

Having in mind from section two how Hume and Smith distance themselves from self-love, and from section three how one may conceive of sympathy, this fourth section provides a simplified formalization of individual preferences. The formalization is later used as a building block in the description of Hume's political society.

In Hobbesian society man is egoistic, but not solely in the sense that only one's own wellbeing matters, whereas that of others is irrelevant. On the contrary, their wellbeing is crucial to our agent. For Hobbes life is like a race, where leaving someone behind is glory. Man seeks something to an eminent degree. Hobbes (1651, ch. VIII) writes:

Vertue generally, in all sorts of subjects, is somewhat that is valued for eminence; and consisteth in comparison.

He continues (Ch. XVII): "But man, whose Joy consisteth in comparing himself with other men, can relish nothing but what is eminent." Simplify and assume two players having payoffs p_1 and p_2 . Each player is egoistic and seeks to maximize eminence over the other player. Using Taylor's (1987, p. 112) formalization, each player seeks to maximize a utility function being a weighted sum of p_1 and p_2 :

$$u_i = a_i p_i + b_i p_j, \quad i, j = 1, 2, \quad i \neq j. \quad (1)$$

Being egoistic means that $a_i > 0$, and seeking eminence over the other player implies that $b_i < 0$. If $b_i < 0$, then player i has malevolence, negative sympathy etc., for player j , denoted as negative altruism. Hence a Hobbesian is egoistic and negatively altruistic.

In our formalization a pure egoist has $b_i = 0$, a pure altruist $a_i = 0$; $b_i < 0$ corresponds to negative altruism, and $b_i > 0$ to positive altruism. If $a_i > 0$, a person is egoistic; if $a_i < 0$, a person is masochistic, non-competitive, ascetic, etc.

Now consider man in Hume's society. Hume assumes self-interest in his political theory: thus $a_i > 0$. But desire for eminence is not predominant in Hume's political theory. Rather benevolence or positive altruism is crucial: thus we set $b_i > 0$, though we do not rule out that $b_i < 0$ also in Hume's society since he also allows resentment to be implanted in human nature, as we will see in the following pages. Hume (1740, p. 367) writes: "The passions of love and hatred are always followed by, or rather conjoin'd with benevolence and anger." Hume continues that there are "certain instincts originally implanted in our natures, such as benevolence and resentment" (p. 417). Denote this benevolence as private benevolence since it is instinctive, and in some sense arises from love. But Hume also assumes the presence of what may be called extensive benevolence. He writes (1740, p. 369): "Pity is a concern for, and malice a joy in the misery of others, without any friendship or emnity to occasion this concern or joy". Hume (1740, p. 579) assumes that extensive benevolence arises from sympathy: "We have no such extensive concern for society but from sympathy". Consonant with the argument in section three, Hume considers sympathy as our ability to "receive by communication... [others'] inclinations and sentiments, however different from, or even contrary to our own". Taylor (1987, p. 152) argues that extensive benevolence is not instinctive, but due to sympathy. This is controversial. So consider Hume (1740, p. 481):

In general, it may be affirm'd, that there is no such passion in human minds, as the love of mankind, merely as such, independent of personal qualities, of services, or of relation to ourself. 'Tis true, there is no human, and indeed no sensible, creature, whose happiness or misery does not, in some measure, affect us, when brought near to us, and represented in lively colours: But this proceeds merely from sympathy, and is no proof of such an universal affection to mankind, since this concern extends itself beyond our own species. An affection betwixt the sexes is a passion evidently implanted in human nature; and this passion not only appears in its peculiar symptoms, but also in inflaming every other principle of affection, and raising a stronger love from beauty, wit, kindness, than wou'd otherwise flow from them. Were there an universal love among all human creatures, it wou'd appear in the same manner.

Hume says that extensive benevolence is due to sympathy which is, "the conversion of an idea into an impression by the force of imagination" (1740, p. 427). In a sense there are no instincts involved here. But in another sense Hume argues that universal love arises, in the last instance, from "affection betwixt the sexes", which is instinctive. In so far as extensive benevolence and universal love are correlated and/or interdependent, Taylor's argument is problematic. Hume is controversial on this point. Not only does he not define his concepts properly, but he seems to have different opinions in his *Treatise* (1740) and in the later *Enquiry* (1751). Hume was reluctant to label

extensive benevolence as an instinct in the former, whereas in the later work, he seems to have considered extensive benevolence to be an instinct. Perhaps Hume felt it necessary to make a more efficient attack on self-love, or perhaps he grew more mature and sincerely felt extensive benevolence to be instinctively actuating.

Relative weighting of self-interest and sympathy in Hume's theory

Hume (1740, p. 487) writes: "I am of opinion, that tho' it be rare to meet with one, who loves any single person better than himself; yet 'tis rare to meet with one, in whom all the kind affections, taken together, do not overbalance all the selfish." Hume says here that most people love themselves more than any other single person, but that the aggregate of a person's concerns for others in most cases overbalances concern for oneself. Hume's argument for this last part is, for example, that a man bestows resources on his wife and children. This may be so, but the problem here is to agree on a scale or method of how to measure concern for oneself relative to concern for others. Some kind of empirical research is necessary to verify Hume here. Hume (1740, p. 481; 1751, Appendix II) assigns high relevance and weight to benevolence. In his *History of England* Hume writes: "the milder views of gratitude, honour, friendship, generosity, are frequently able, among princes as well as private persons, to counterbalance these selfish considerations" (Hirschman, 1977, p. 48).

On the other hand, when Hume explains the origin of justice, property, government, etc. he assumes that benevolence is limited. He writes (1740, p. 534): "Nothing is more certain, than that men are, in a great measure, govern'd by interest, and that even when they extend their concern beyond themselves, 'tis not to any great distance; nor is it usual for them, in common life, to look farther than their nearest friends and acquaintance." He continues (1740, p. 488): "our strongest attention is confin'd to ourselves; our next is extended to our relations and acquaintance; and 'tis only the weakest which reaches to strangers and in different persons." Smith's (1776, part I.ii, 2) well-known view is related: "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." Thus self-interest is crucial in Hume's as well as Smith's theory, as section six, following, proceeds to show.

Hume's conventions

Any modern economic textbook equips persons with self-interest or opportunism and bounded rationality, and the world with scarcity, complexity and uncertainty. This terminology is consonant with Hume's. With regard to property, Hume's view is that "the principal disturbance in society arises from those goods which we call external" (1740, p. 489). The first reason for this is "the instability of their possession, along with their scarcity" (p. 488). The second is that men are self-interested (p. 491):

This avidity alone, of acquiring goods and possessions for ourselves and our nearest friends, is insatiable, perpetual, universal, and directly destructive of society. There scarce is any one, who is not actuated by it; and there is no one, who has not reason to fear from it, when it acts without restraint... 'Tis certain, that no affection of the human mind has both a sufficient force, and a proper direction to counter-balance the love of gain.

Consequently, Hume's state of nature is not too different from Hobbes' with war of all against all. Hume does not exactly consider "the life of man, [as] solitary, poore, nasty, brutish, and short" (Hobbes, 1651, Ch. XIII), but he clearly perceives the influence of benevolence as weak, in his political theory. He says (Hume, 1740, p. 492): "There is no passion, therefore, capable of controlling the interested affection, but the very affection itself, by an alteration of its direction." To solve this problem, Hume continues (p. 489): "The remedy, then, is not deriv'd from nature, but from artifice... This can be done after no other manner, than by a convention enter'd into by all members of the society to bestow stability on the possession of those external goods, and leave every one in the peaceable enjoyment of what he may acquire by his fortune and industry." Complete stability of possession is inconvenient, so Hume argues for a second convention, that of "transference of property by consent" (1740, p. 514). Also a third convention is necessary, that of "the performance of promises" (1740, pp. 516 and 526). (For a discussion of the matter of convention the reader might be interested in considering Lewis, 1969.)

Scrutinize more closely what Hume means by a *convention*. His definition (1740, p. 490) is as follows:

convention is not of the nature of a promise: For even promises... arise from human conventions. It is only a general sense of common interest; which sense all the members of the society express to one another, and which induces them to regulate their conduct by certain rules. I observe, that it will be for my interest to leave another in the possession of his goods, provided he will act in the same manner with regard to me. He is sensible of a like interest in the regulation of his conduct. When this common sense of interest is mutually express'd, and is known to both, it produces a suitable resolution and behaviour. And this may properly enough be call'd a convention or agreement betwixt us, tho' without the interposition of a promise; since the actions of each of us have a reference to those of the other, and are perform'd upon the supposition, that something is to be perform'd on the other part. Two men, who pull the oars of a boat, do it by an agreement or convention, tho' they have never given promises to each other. Nor is the rule concerning the stability of possession the less deriv'd from human conventions, that it arises gradually, and acquires force by a slow progression, and by our repeated experience of the inconveniences of transgressing it.

Taylor (1987, p. 157), in Hume's spirit, defines a convention as: "A regularity R in the behaviour of members of a population P when they are agents in a recurrent situation S is a convention if and only if, in any instance of S among members of P, (1) everyone conforms to R; (2) everyone expects everyone else to conform to R; (3) everyone prefers to conform to R on condition that the others do, since S is a co-ordination problem and uniform conformity to R is a co-ordination equilibrium in S." A co-ordination equilibrium is here defined as a strategy vector such that no player can obtain a larger payoff if he or any other player unilaterally uses a different strategy. Thus a co-ordination equilibrium is a Nash equilibrium, but not conversely.

Consider as illustration the following co-ordination problem known as the "battle of the sexes" game, shown in Figure 1.

A married couple attempts to agree whether to spend the evening at the cinema (C) or at the theatre (T). Neither wants to spend the evening alone,

thus strategy pairs (T,C) and (C,T) (for husband, wife) give a payoff of (0,0). The husband prefers the cinema and the culturally-inclined wife the theatre, with payoffs of (2,1) or (1,2), corresponding to the two pure strategy Nash equilibria in this game. In games like this, anything that tends to focus the players' attention on one equilibrium rather than another may make them both expect it and hence fulfil it, like a self-fulfilling prophecy. This is called the focal-point effect (Schelling, 1960). If our couple lives in a society where wives traditionally defer to their husbands, one possible resolution of the game is to agree on the cinema, giving a payoff of (2,1) to husband, wife. Even if the couple feels no compulsion to conform to this tradition, the wife may potentially expect that her husband will presume that he should choose the cinema, so that she will reluctantly agree on the cinema, and the husband may expect his wife to choose the cinema, implying that the cinema is preferable for them both.

		Wife	
		C	T
Husband	C	2, 1	0, 0
	T	0, 0	1, 2

Figure 1.
The battle of the sexes

An alternative co-ordination problem is whether to drive on the right or the left side of the road, which corresponds to the payoffs (1,1), (0,0), (0,0), (1,1) for strategies (R,R), (L,R), (R,L), (L,L), where R = right and L = left. In this game no one has a specific preference regarding which side she drives on as long as everybody else drives on the same side as she does. In Hume's definition, conventions emerge spontaneously. But Hume's conventions also persist spontaneously, as with the two men who pull the oars of a boat in order to cross a lake. The reason for this is that a convention is an equilibrium from which no one has an incentive to deviate unilaterally. If one of the men in the boat stops rowing, the boat will start going in a circle. This may correspond to this man's preferences for a while, but given that he seeks to cross the lake, sooner or later his incentives will tell him, by his own free will, voluntarily, to continue rowing.

A crucial insight here is that the way Hume defines *convention* (1740, p. 490), no government or external agency is required to constrain or compel people to conform. Again, this is because a convention is an equilibrium from which no one has any incentive to deviate unilaterally. But, clearly, Hume proceeds to argue for the necessity of government to constrain men, and to ensure adherence to laws of justice. Necessity of government is because people find themselves not in recurrent co-ordination games, but rather, as shown in the subsequent sections, in recurrent or iterated games having other characteristics.

Defining convention as a co-ordination equilibrium leads Taylor (1987, p. 158) to argue that the laws of justice cannot be conventions in the way Hume defines conventions since men adhere voluntarily to conventions but

not necessarily to laws of justice. Taylor's argument is that it "would be proper to call the laws of justice 'conventions' only if all men preferred any system of such laws (and therefore any distribution of possessions) to no laws at all and were indifferent (or nearly so) between all possible systems." This first condition is accepted by Hume, but the second is not. The reason is that people are certainly not indifferent to different distributions of property, and hence neither are they indifferent to different laws of justice determining distributions. Clearly, government is needed. Before introducing government, let us consider various types of games describing political societies.

The assurance game, chicken game and the prisoner's dilemma

This seventh section confines attention to games with two players. The assurance game is given in Figure 2.

This game has two Nash equilibria: (C,C) and (D,D) giving payoffs (7,7) and (2,2). In a sense this is the co-ordination game which we described in the last section. However, since both players prefer (C,C) to (D,D), no one will expect the latter, implying (C,C) as a unique Pareto-optimal outcome. An example of this game is the production of a public good where no amount of the good can be provided until total contributions exceed some threshold. This is frequently the case with bridges, roads, etc. If no-one contributes to the good, assume they both get a payoff of 2. If player 1 contributes to the good by choosing C thereby having a cost of 2, whereas player 2 does not contribute by choosing D, the payoff is (0,2) since the public good is not produced. However, if player 2 also decides to contribute, by choosing C and having a cost of 2, the good will be produced, for their mutual benefit, and a payoff of, for example, (7,7).

Figure 2.
The 2×2 assurance
game

	C	D
C	7, 7	0, 2
D	2, 0	2, 2

Clearly, situations are not always this beneficial. Consider the chicken game, shown in Figure 3.

This game has two Nash equilibria in pure strategies: (D,C) and (C,D) giving payoffs of (7,2) and (2,7), whereas (C,C) giving (6,6) is collectively optimal. The difference from the assurance game is that each player in the chicken game can produce some of the public good alone by choosing C, and will indeed choose C if the other player chooses D, thereby getting a payoff

Figure 3.
The 2×2 chicken
game

	C	D
C	6, 6	2, 7
D	7, 2	0, 0

of 2 rather than 0. But, more crucial, each player prefers D given that the opponent chooses C; that is, each player prefers to free ride on the other's contribution. Thus there is a pre-commitment advantage. If one player is able to pre-commit to defect, the other player is "forced" into co-operation. One example (Taylor, 1987, p. 37) of chicken preferences is two factories discharging polluted material into a small lake. If both discharge (D,D), the pollution exceeds a critical threshold implying an ecological catastrophe so bad that each factory prefers unilaterally to refrain from pollution if the other pollutes. Thus (C,D) is preferred over (D,D). Nevertheless, each factory prefers a free ride on the restraint of the other factory (D,C), to mutual co-operation (C,C).

An alternative example of a chicken game is two teenagers driving against each other, each in his own car. D involves driving straight ahead, and C involves making a turn to one or the other side. If both drive straight ahead, (D,D), a collision will ensue, with likely death and a payoff of (0,0) (or $-1,000, -1,000$). If one drives straight ahead and the other makes a turn (D,C), the first gets a payoff of 7 because he preserves his life as well as obtaining a reputation for cold-bloodedness in hazardous situations, whereas the other gets a payoff of 2. This last payoff is larger than 0 because death is avoided, but smaller than 6 since a bad reputation, for cowardice, ensues. If both make a turn (C,C), payoffs are (6,6) since no deaths occur, and since there are neither good nor bad reputational effects. How does one get out of a chicken game? Consider Figure 4. It can be shown that there exists a mixed strategy equilibrium ($2/3C, 2/3C$) giving payoffs of ($14/3, 14/3$) where each player chooses C with probability ($2/3$) and D with the remaining probability $1/3$. Thus there are three equilibria in our chicken game, and no better solution than these three exists, given that our players get no outside help.

On rational grounds there is no way in which the players can choose between these equilibria. In a sense the mixed strategy stands out as a focal point, but a problem is that there usually exists doubt about the psychology of the other player. A crucial feature of this chicken game is that our players may do better by getting external help, for example, from a mediator. The players and the mediator may collectively agree on a randomization device which is common knowledge among the players. Assume they agree that (D,D) with a payoff of (0,0) is to be avoided, that both players are to get the same expected payoff since the payoff-structure is symmetrical, and that this expected payoff is to be maximized. Moreover, require that the mediator's recommendation is such that both players have incentives to abide. That is, their expected payoff will be lower if they do not follow the mediator's recommendation, no matter what the mediator might recommend. Game-theoretically it can be shown that our players can obtain the expected payoff of ($21/4, 21/4$) by choosing (C,C) with probability ($1/2$), (D,C) with probability ($1/4$), and (C,D) with the remaining probability ($1/4$), as determined by the mediator by, for instance, flipping a coin, then subsequently and secretly informing the players, separately, of the outcome of the coin toss.

Although assurance and chicken games are frequently descriptive, the best-known game is the prisoner's dilemma (PD), shown in Figure 4.

This game has one equilibrium in dominant strategies (D,D) with a payoff of (1,1). No matter what the other player does, defect (D) is rational. Hobbes considers society as a PD where (D,D) corresponds to war, and (C,C) to peace. Hobbes' first law of nature is a suggestion for getting out of the mutually destructive (D,D): "That every man, ought to endeavour Peace, as farre as he has hope of attaining it; and when he cannot obtain it, that he may seek, and use, all helps, and advantages of Warre" (1651, Ch. XIV). Hobbes here says that D is rational if one thinks that one cannot obtain peace; that is, if one thinks that some others will choose D. Hume describes

Figure 4.
The 2 × 2 prisoner's
dilemma

	C	D
C	2, 2	0, 3
D	3, 0	1, 1

a hold-up problem, which we with suitable arrangements of payoffs may consider as a PD, as follows (1740, p. 520):

Your corn is ripe today; mine will be so tomorrow. 'Tis profitable for us both, that I should labour with you to-day, and that you should aid me to-morrow. I have no kindness for you, and you have as little for me. I will not, therefore, take any pains upon your account; and should I labour with you upon my own account, in expectation of a return, I know I shou'd be disappointed, and that I shou'd in vain depend upon your gratitude. Here then I leave you to labour alone: You treat me in the same manner. The seasons change; and both of us lose our harvests for want of mutual confidence and security.

Hume's (1740, p. 521) solution to this is to suggest that "we can better satisfy our appetites in an oblique and artificial manner, than by their headlong and impetuous motion. Hence I learn to do a service to another, without bearing him any real kindness; because I forsee, that he will return my service, in expectation of another of the same kind, and in order to maintain the same correspondence of good offices with me or with others." Hume writes this in his section about the obligation of promises, and he continues: "After these signs [of promise] are instituted, whoever uses them is immediately bound by his interest to execute his engagements, and must never expect to be trusted any more, if he refuses to perform what he promis'd" (1740, p. 522).

One of the primary functions of government is getting out of PD situations. And one effective way of doing this is by ensuring adherence to Hume's three laws of nature, those of the stability of possession, its transference by consent and the performance of promises.

Hume's political game

Static game (section eight)

Confine attention to two players. Assume a symmetrical game with a payoff structure as shown in Figure 5.

- This game is an assurance game if
 $x > y$ and $w > z$; (1)
- a chicken game if
 $y > x$ and $z > w$; (2)
- and a PD if
 $y > x > w > z$. (3)

Elster (1982), for example, assumes that $y > w$ for the assurance game and $x > z$ for the chicken game. These requirements are not crucial to our argument, so they are not included. As we saw in section four, in both Hobbes' and Hume's society not only one's own payoff, but also the other player's payoff are relevant for one's own total utility. Using equation (4.1 (from section 4)), the utility structure, corresponding to the payoff structure in Figure 5, is as in Figure 6 (Taylor, 1987, p. 113).

- This game is an assurance game if
 $a_i y + b_i z < (a_i + b_i) x \Rightarrow a_i < b_i (x - z)/(y - x)$, (4)
- and if
 $(a_i + b_i) w > a_i z + b_i y \Rightarrow a_i > b_i (y - w)/(w - z)$. (5)
- The game is a chicken game if
 $a_i y + b_i z > (a_i + b_i) x \Rightarrow a_i > b_i (x - z)/(y - x)$, (6)
- and if
 $(a_i + b_i) w < a_i z + b_i y \Rightarrow a_i < b_i (y - w)/(w - z)$. (7)
- And the game is a PD if $(x - z) > (y - x)$,
 $a_i y + b_i z > (a_i + b_i) x \Rightarrow a_i > b_i (y - w)/(w - z)$, (8)
- if
 $(a_i + b_i) w > a_i z + b_i y \Rightarrow a_i > b_i (y - w)/(w - z)$, (9)
- and if
 $(a_i + b_i) x > (a_i + b_i) w \Rightarrow a_i + b_i > 0$. (10)

	C	D
C	x, x	z, y
D	y, z	w, w

Figure 5.
Pay-off structure in
 2×2 symmetric
game

	C	D
C	$(a_1 + b_1)x, (a_2 + b_2)x$	$a_1 z + b_1 y, a_2 y + b_2 z$
D	$a_1 y + b_1 z, a_2 z + b_2 y$	$(a_1 + b_1)w, (a_2 + b_2)w$

Figure 6.
Utility structure in
 2×2 symmetrical
game

Assume a Hobbesian society and a PD where equation (3) is satisfied. Section four showed that a Hobbesian seeks to maximize a weighted sum of egoism and negative altruism; i.e.

$$a_i > 0, b_i < 0, i = 1, 2. \quad (11)$$

Observe that if (3) and (11) are satisfied, then (8) and (9) are easily satisfied, but (10) not necessarily so. However, given that (10) is satisfied, the transformed game corresponding to the utility structure in Figure 6 is a more severe PD than the untransformed game corresponding to the payoff structure in Figure 5. Hence negative altruism due to seeking eminence over one's opponent has the negative implication of rendering an already undesirable PD into a worse PD. Alternatively, if the Hobbesian society is a chicken game and (11) is satisfied, observe that b_i negative may imply that (7) may or may not be satisfied. Hence the transformed game may be a chicken game or a PD.

Consider a PD in Hume's spirit where (3) is satisfied. Section four showed that a person in Hume's society seeks to maximize a weighted sum of egoism and positive altruism:

$$a_i > 0, b_i > 0, i = 1, 2. \quad (12)$$

If (3) and (12) are satisfied, then (10) is satisfied, but (8) and (9) are not necessarily so. Hence the transformed game may or may not be a PD. In fact, the transformed game may be a chicken game, an assurance game, or perhaps even a game in which C dominates D for both players. Hence positive altruism due to having benevolence for one's opponent may have the positive implication of rendering an undesirable PD a more beneficial PD, or, at best, into a game where mutual co-operation is guaranteed. Alternatively, if Hume's society is a chicken game and (13) is satisfied, we see that b_i positive cannot have that influence on (8) that the transformed game is rendered into a PD, as we have seen may be possible with negative altruism in a Hobbesian society. The transformed game may be a chicken game, or a game where prospects for co-operation are better.

Dynamic game

Time does not play an explicit role in Hobbes' theory, and one interpretation involves considering his society as a one-shot PD. This view is consonant with most of Hobbes' work. But let us state Hobbes' (1651, Ch. XIV) second law of nature, following his first presented in the previous section: "That a man be willing, when others are so too, as farre-forth, as for Peace, and defence of himself he shall think it necessary, to lay down this right to all things; and be contented with so much liberty against other men, as he would allow other men against himself." Hobbes here says that it is rational for a player to choose C – lay down one's right of nature – if other players do. This statement is of course wrong if one is situated in a PD, since D then is a dominant strategy. Hobbes is correct if government has been able to alter the game characteristics so that the game is no longer a PD, but rather a game where (C,C) is an equilibrium. A second interpretation of Hobbes may arise from his statement "when others are so too" in the second law of nature, and involves considering his society as an iterated PD, with a

possibility for a conditional choice of action, that is, choosing C if other players have chosen C in the previous round. The problem with this interpretation is that Hobbes has little to say about dynamic interaction among players.

Hume, on the other hand, makes throughout his works explicit elaborations of the time aspect, and his crucial point is that we discount future benefits. Hume has a section (1740, p. 428) about “distance in space and time”, where he says that “men are principally concern’d about those objects, which are not much remov’d either in space or time, enjoying the present, and leaving what is afar off to the care of chance and fortune”. He maintains that “the consequence of a removal in space are much inferior to those of a removal in time” (1740, p. 429). We may argue that some discounting of the future is rational and part of sound judgement, due to risk, uncertainty, inflation, transaction costs, etc. Hume’s view, however, is that people discount the future far beyond what maximizing long-range utility would teach. Hume (1740, p. 538) holds that “there is no quality in human nature, which causes more fatal errors in our conduct, than that which leads us to prefer whatever is present to the distant and remote, and makes us desire objects more according to their situation than their intrinsic value”. He argues that men “prefer any trivial advantage, that is present, to the maintenance of order in society” (1740, p. 538; see also pp. 499, 539 and 545 for further comments regarding time).

Consider a game-theoretic formalization of Hume’s theory. As argued at the start of this section, the prospects for co-operation are greater in Hume’s society than in Hobbes’ because of Hume’s assumption of positive rather than negative altruism in human nature. Thus if society is a moderately severe PD, then Hume’s PD is less severe than Hobbes’. Nevertheless, one is still situated in a PD, and in a PD non co-operation is the dominant strategy. That is, D is rational regardless of others’ behaviour. A better description of the human condition involves considering iterative games. Iteration increases prospects for co-operation, since the players can condition their strategies on what has happened in previous rounds. One example of work in this area is Axelrod’s. Axelrod (1984) shows how co-operation, based on tit-for-tat strategy, can get an initial foothold in an egoistic world, can thrive while interacting with other strategies, and can resist invasion once fully established. Intertemporal linkages between different rounds in a game give opportunities for the players to nurture various kinds of reputation, such as for co-operative behaviour. It is crucial to observe that whereas iteration of a PD increases prospects for co-operation, discounting of future benefits reduces such prospects. Thus these two forces, iteration and discounting, pull in opposite directions.

There is a third factor reducing prospects for co-operation, that of hidden action. It is frequently the case that no one, except the actor herself, can observe whether or not she co-operated. Thus the possibility of detection of deceitful actions, D, is imperative. Formalize this in an infinitely repeated PD with two players, using payoffs from Figure 4. Mutual co-operation (C,C) gives a payoff of (2,2), mutual deception (D,D) (1,1), free-riding (D,C) 3, whereas being exploited (C,D) gives 0. Detection of deception (see Frank, 1988) implies mutual deception for the remainder of the game. Assume

discount factor δ , and that deception is detected with probability q . Mutual co-operation yields 2 in every period with an expected payoff of $2/(1 - \delta)$ (geometric series). Unilateral deception in the first period gives 3, and then 1 or 2 thereafter. Thus, detection of deception in the first period yields an expected payoff of $1/(1 - \delta)$ with probability q , and $2/(1 - \delta)$ with probability $1 - q$, considered from period 2 and thereafter. Hence co-operation is preferred if $2/(1 - \delta) > \delta[q/(1 - p) + (1 - q)2/(1 - \delta)]$, i.e. if $\delta > 1/(1 + q)$. Observe that if there is a 100 per cent chance of detection of deception ($q = 1$), co-operation is rational if the discount factor is $1/2$ or higher. With a 50 per cent chance of detection, a discount factor of $2/3$ is necessary to rationalize co-operation. In other words, if detection of deception is very unlikely, a higher emphasis on the future is necessary to rationalize co-operation. Features of this kind are crucial in the folk theorem (Fudenberg and Maskin, 1986) which states that any individually rational payoff vector in a one-shot game of complete information can arise in a perfect equilibrium of the corresponding infinitely repeated game if the players discount the future sufficiently little (δ is high). Hume, and Axelrod (1984), agree on a remedy for this inconvenience: one is to render something distant (the future) the immediate interest of the actor (Hume, 1740, p. 537). Axelrod (1984, p. 126) recommends enlarging the shadow of the future by making interactions more durable and frequent. The final section about government considers Hume's remedies. But first we must introduce a game with more than two players.

Game with more than two players

We have seen that Hobbes considers society as a static PD. Moreover, Hobbes does not discuss societies with only a few players, but confines attention to many players, apparently corresponding to large societies like the one in which he lived. Hume discusses both small and large societies, and is aware that a small society which is not a PD may be turned into one by adding more players. Hume further considers man as situated in an infinitely repeated game, presumably a PD, with discounting (and perhaps hidden action and hidden information), and with many players. We may interact more with some players than with others, and we may not know with whom we are interacting. This reduces prospects for co-operation. Hume's best description of the problem with many players is his "drained meadow" example:

Two neighbours may agree to drain a meadow, which they possess in common; because 'tis easy for them to know each other's mind; and each must perceive, that the immediate consequence of his failing in his part, is, the abandoning the whole project. But 'tis very difficult, and indeed impossible, that a thousand persons shou'd agree in any such action; it being difficult for them to concert so complicated a design, and still more difficult for them to execute it; while each seeks a pretext to free himself of the trouble and expence, and wou'd lay the whole burden on others (1740, p. 538).

Hume here concentrates on a static situation, and clearly perceives the two neighbours as not being placed in a PD. Compare this example with the harvest example (Hume, 1740, p. 520) described in the previous section, which clearly is a PD. The difference between the examples is that our

neighbours own the meadow in common and may drain it simultaneously. The farmers, however, have each their own harvest, and co-operation requires adherence to an artificially constructed promise where effort at separate points in time is involved. With 1,000 players, however, Hume considers the “drained meadow” example as a PD, and he gives two arguments why large societies fail to provide public goods like a drained meadow. The first is that each player has an incentive to free ride on the others’ contribution. The second is that larger groups have higher costs of organization.

We consequently see that although iteration of a game may increase prospects for co-operation, discounting of the future, many players and hidden action and information reduce such prospects. Following up Hume’s “discount examples” from earlier, Hume (1740, p. 535) continues:

The consequences of every breach of equity seem to lie very remote, and are not able to counterbalance any immediate advantage, that may be reap’d from it. They are, however, never the less real for being remote; and as all men are, in some degree, subject to the same weakness, it necessarily happens, that the violations of equity must become very frequent in society, and the commerce of men, by that means, be render’d very dangerous and uncertain. You have the same propension, that I have, in favour of what is contiguous above what is remote. You are, therefore, naturally carried to commit acts of injustice as well as me. Your example both pushes me forward in this way by imitation, and also affords me a new reason for any breach of equity, by shewing me, that I shou’d be the cully of my integrity, if I alone shou’d impose on myself a severe restraint amidst the licentiousness of others.

Hume’s remedy for this is to introduce government.

Government

(Section nine)

Hume (1740, p. 535) writes: “The remedy can only come from the consent of men; and if men be incapable of themselves to prefer remote to contiguous, they will never consent to any thing, which wou’d oblige them to such a choice, and contradict, in so sensible a manner, their natural principles and propensities.” It is imperative to observe here that Hume’s key argument for the necessity of government is, in fact, solidly founded on the discount factor. He argues (1740, p. 537):

The only difficulty, therefore, is to find out this expedient, by which men cure their natural weakness, and lay themselves under the necessity of observing the laws of justice and equity, notwithstanding their violent propension to prefer contiguous to remote. ’Tis evident such a remedy can never be effectual without correcting this propensity; and that ’tis impossible to change or correct any thing material in our nature, the utmost we can do is to change our circumstances and situation, and render the observance of the laws of justice our nearest interest, and their violation our most remote.

Hume (1740) starts his section about the “origin of government” by arguing that men are “govern’d by interest”. It is likely that with the words “material in our nature” in the previous passage, in addition to the discount factor, Hume has our propensity to self-interest in mind. Hume’s remedy is to render something remote one’s near interest, or in Axelrod’s (1984) words,

to enlarge men's shadow of the future. This can only take place in respect of a few: civil magistrates, kings, ministers, etc., "who being indifferent persons to the greatest part of the state, have no interest or, but a remote one, in any act of injustice; and being satisfied with their present condition, and with their part in society, have an immediate interest in every execution of justice, which is so necessary to the upholding of society" (Hume, 1740, p. 537):

Magistrates find an immediate interest in the interest of any considerable part of their subjects... Thus bridges are built; harbours open'd; ramparts rais'd; canals form'd; fleets equip'd; and armies disciplin'd; every where, by the care of government, which, tho' compos'd of men subject to all human infirmities, becomes, by one of the finest and most subtle inventions imaginable, a composition, that is, in some measure, exempted from all these infirmities (Hume, 1740, pp. 538-39).

Thus, self-interest and the discounting of the future are key ingredients in Hume's political theory, and are the main arguments for the necessity of government. However, having in mind the influence of factors such as sympathy and benevolence in Hume's theory, considered in earlier sections, let us conclude this section by observing the following passages which illustrate how Hume considers self-interest and sympathy to interact in his political theory: He says (1740, p. 499): "Thus self-interest is the original motive to the establishment of justice: but a sympathy with public interest is the source of the moral approbation, which attends that virtue." Hume (1740, p. 577) holds that sympathy "produces our sentiment of morals in all the artificial virtues". In a sense Hume's view is that self-interest is key factor, whereas sympathy comes afterwards and gives rise to moral approbation. Hume (p. 579) continues: "After... [the whole scheme of the laws of justice] is once establish'd by these conventions, it is naturally attended with a strong sentiment of morals; which can proceed from nothing but our sympathy with the interests of society." Thus Hume holds that even though justice is an artificial construct, we have a natural inclination morally to approve. We naturally approve of artificially constructed systems, that these systems may be a result of historical evolution and hence to some extent subject to fortune; and accidents do not, in Hume's view, alter men's propensity naturally to approve. Hume (p. 619) says:

Tho' justice be artificial, the sense of its morality is natural. 'Tis the combination of men, in a system of conduct, which renders any act of justice beneficial to society. But when once it has that tendency, we naturally approve of it; and if we did not so, 'tis impossible any combination or convention cou'd ever produce that sentiment.

Conclusion

We have analysed the impact of self-interest and sympathy in Hume's theory. Hume considers sympathy a powerful motivator, which takes us so far out of ourselves, as to give us the same pleasure or uneasiness in the characters of others as if they had a tendency to our own advantage or loss. Smith elaborates on our specific ability to enter into the particular situation in which others are positioned, thereby sharing by sympathy (e.g. the

gratitude evoked by others' benevolence). Hume and Smith also consider self-interest as actuating, especially in economic affairs. Section four formalized a utility function embodying self-interest and sympathy. I showed that a Hobbesian maximizes a weighted sum of egoism and eminence (negative altruism), whereas people in Hume's society maximizes a weighted sum of egoism and benevolence (positive altruism). The fifth section assessed the relative weight of self-interest and sympathy in Hume's political theory, admitting that Hume considered benevolence as limited when it comes to explaining the origin of justice, property and government. The next section defined Hume's concept of convention as a general sense of common interest, which sense all members of a society express to one another, and which induces them to regulate their conduct by certain rules. I described two co-ordination problems, the battle of the sexes, and whether to drive on the right or the left side of the road. The focal-point effect is illustrated, potentially enabling us to choose between equally beneficial equilibria. Having defined conventions as co-ordination equilibria, from which no one has any incentives unilaterally to deviate, I disputed Hume's view that laws of justice are conventions. Admitting that political societies are not necessarily co-ordination games, section seven illustrated features of three typical games; the assurance game, the chicken game, and the prisoner's dilemma. Political societies may be described by these games, or by hybrids of them.

Section eight used the formulation of section four to distinguish Hobbes' from Hume's society. It showed that if a political society is best described with payoffs, as in the prisoner's dilemma, Hobbes' assumptions lead people to play a more severe prisoner's dilemma than is the case with Hume's assumptions. At best, assumptions of benevolence may lead people to play a game which is no longer that of the prisoner's dilemma. Further, I showed how a chicken game can be transformed into a prisoner's dilemma with Hobbes' assumptions about eminence. This is not possible with Hume's assumptions; at the very best benevolence may turn a chicken game into a game where co-operation is a dominant strategy for all players. Discussion then elaborated on the time aspect which plays such a crucial role in Hume's, but not in Hobbes', theory. We saw that although iteration of the prisoner's dilemma increases prospects for co-operation, discounting future benefits reduces such prospects. I illustrated this with an infinitely repeated two-person prisoner's dilemma with discounting, including the possibility that a player imperfectly observes the opponent's action in the preceding round. Detection of whether the opponent had co-operated rather than defected in the preceding round was done with a probability possibly of less than one. I concluded that reduced possibilities for detection reduce prospects for co-operation. The section ended by introducing more than two players. Hume's "drained meadow" example was used to show that 1,000 neighbours, but not necessarily two, are likely to be situated in a prisoner's dilemma. Hume's explanation is consonant with public goods theory: in large societies each player has incentives to free ride on others' contributions, and large groups have higher costs of organization.

Hume's key arguments for the necessity of government are that men are self-interested and discount future benefits beyond what sound, long-range

judgement teaches; that is, men prefer any trivial advantage that is present, especially when it is brought near and represented in lively colours, to the maintenance of order in society. Having thus observed the impossibility of correcting anything material in human nature, Hume concludes that the utmost we can do is to change our circumstances and situation, by instituting government to render the observance of the laws of justice our nearest interest, and their violation our most remote interest. Finally, once the laws of justice are established, they are naturally attended by a strong sentiment of morality, which can proceed from nothing but our sympathy with the interests of society.

References

- Axelrod, R. (1984), *The Evolution of Cooperation*, Basic Books, New Yorks, NY.
- Elster, J. (1982), *Explaining Technical Change*, Cambridge University Press, Cambridge.
- Elster, J. (1983), *Sour Grapes*, Cambridge University Press, Cambridge.
- Elster, J. (1989), *The Cement of Society*, Cambridge University Press, Cambridge.
- Frank, R.H. (1988), *Passions within Reason*, Norton & Co., New York, NY.
- Fudenberg, D. and Maskin, E. (1986), "The folk theorem in two-person repeated games with discounting and incomplete information", *Econometrica*, Vol. 54, pp. 533-54.
- Gauthier, D.P. (1986), *Morals by Agreement*, Oxford University Press, Oxford.
- Hardin, R. (1988), *Morality within the Limits of Reason*, Chicago University Press, Chicago, IL.
- Hirschman, A.O. (1977), *The Passions and the Interests*, Princeton University Press, Princeton, NJ.
- Hobbes, T. (1651), *Leviathan*, J.M. Dent & Sons Ltd, London (1973).
- Hume, D. (1740), *A Treatise of Human Nature*, Selby-Bigge, L.A. (Ed.), Clarendon Press, Oxford, (1978).
- Hume, D. (1751), *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, Selby-Bigge, L.A. (Ed.), 3rd ed. (1975), Clarendon Press, Oxford.
- Hume, D. (1777), *Essays, Moral, Political, and Literary*, Miller, E.F., Liberty Classics, Indianapolis, IN (1985).
- Hume, D. (1779), *Dialogues Concerning Natural Religion*, Macmillan, New York, NY.
- Lewis, D.K. (1969), *Convention*, Harvard University Press, Cambridge, MA.
- Myers, M.L. (1983), *The Soul of Modern Economic Man, Ideas of Self-Interest, Thomas Hobbes to Adam Smith*, Chicago University Press, Chicago, IL.
- Nietzsche, F. (1886), *Beyond Good and Evil*, Vintage Books, New York, NY (1966).
- Perrow, C. (1986), *Complex Organizations*, Random House, New York, NY.
- Schelling, T.C. (1960), *The Strategy of Conflict*, Harvard University Press, Cambridge, MA.
- Sen, A. (1987), *On Ethics and Economics*, Blackwell, Oxford.
- Smith, A. (1759), *The Theory of Moral Sentiments*, Oxford University Press, Oxford (1976).
- Smith, A. (1776), *The Wealth of Nations*, Everyman's Library, New York, NY (1910).
- Taylor, M. (1987), *The Possibility of Cooperation*, Cambridge University Press, Cambridge.