Using the economic method to investigate drugs in sport

Keiran Sharpe

School of Business, UNSW@ADFA, Canberra, ACT, 2600

Abstract: This paper explains how the economic method can be applied to explain the phenomenon of the use of drugs in sports. The first part of the paper explains the general character of the economic method; and the second part applies the method to the case at issue.

Introduction

In this paper, we explore the way in which the economic method sheds light on the phenomenon of drug usage in sports contests. The paper has two sections: first, it looks at the economic method itself, and, secondly, it shows how that method is applied to the issue at hand.

I. The economic method

Economics is the study of rational behaviour – i.e. it's the study of the way in which reasoning persons decide to act in pursuit of the aims they have. Three implications of this statement are worth commenting on: first, the statement indicates that economics is a *teleological* method; secondly, economics models *agency*; and third, economics is defined, not by the phenomena it explains, but by the *method* it uses to explain the phenomena it examines. Hence, the statement is, in itself, enough to distinguish economics from the other social sciences.

Let's take the first implication. In economics, people aim to bring about states of the world that they prefer. In supposing this, economics traces its intellectual pedigree back to Aristotle, who founded the teleological tradition of decision-making theory (in the *Nicomachean Ethics*, ch.I). However, in contradistinction to the approach of 'The Philosopher', economics abjures any assessment of the worth or merit of the ends at which persons aim. This is to say that economics provides no analysis of what constitutes the good life for man. Instead, it draws on the modern utilitarian and liberal tradition, initiated by Bentham and John Stuart Mill,

which holds that "push-pin is of equal value with the arts and sciences of music and poetry" (Bentham, 1825, ch.I; Mill, 1863, ch.ii). In other words, for a person to be 'rational' in the economist's sense of the word, it matters not at *what* the person aims, rather what matters is that his actions are directed at attaining what the person *most* prefers. The only conditions on the aims that persons have are that those aims be consistent in certain well-defined and obvious ways (e.g. if I prefer, now, to have a cappuccino to a caffè latte, I cannot simultaneously prefer the reverse; or, if I prefer a cappuccino to a latte, and a latte to a macchiato, then I prefer a cappuccino to a macchiato; etc.).

In adopting a teleological approach, economics is distinct from those other social sciences in which it's assumed that people are in the grip of external forces that they cannot wilfully apprehend – such as is generally supposed to be the case in psychology, sociology and anthropology (which adopt some kind of structural determinism according to their fundamental premises). In being teleological, economics does, however, share the basic methodological presupposition of one of the social sciences, viz., political science. Economics is distinct from political science, though, in supposing that the aims that people pursue are sufficiently well-defined to be able to be formally modelled.

We turn now to the second implication of the economic method, viz., the assumption that persons are agents. This supposition is closely related to that of ends-orientation just mentioned, for it hardly makes sense to suppose that persons aim at ends – and that

their actions can be explained in terms of those aims – if they cannot act in pursuit of the aims they have. 'Aiming implies agency' one might say. Again, this supposition situates economics in the lineage of decision-making theory that goes back to Aristotle. And, as above, this assumption separates economics from most of the other social sciences, for which agency is nugatory (because persons are only the epiphenomena of deeper causal structures, as is supposed in psychology, sociology and anthropology). Again, the social science to which economics is most proximate is political science, but it is distinct from that discipline because economics assumes that 'players' are typically rational and are not in the grip of ideological or other impulses, nor are they assumed to be unwittingly subject to the manipulation of others. In the economic analysis of political behaviour everyone is a Machiavellian, or at least has the rational capacities of The Prince.

Finally, turning to the third implication, we note what is perhaps the most curious feature of economics – viz., that it is defined by the method it uses to explain phenomena rather than by the phenomena it explains. The peculiarity of this feature arises from the fact that this way of marking off the discipline is unique, not only in the social sciences, but in the sciences *tout court*. The economic method is simply that mode of analysis in which behaviour is explained in terms of the rational pursuit of the aims that agents have. Hence, any human or social phenomenon that is explained in those terms is being given an economic explanation. Economics isn't just the analysis of relative prices, interest rates, wage rates, growth rates of GDP, etc., but of

marriage, suicide, how to play tennis and poker, treaty negotiations, the waging of war, and so on.

Recently, of course, the economic method has been imported into sociology and political science (where it goes under the name 'rational choice theory') and so those disciplines now give the appearance, at least in part, of being economics-like in their approach. Conversely, economics has recently imported certain elements of psychology – especially those aspects that concern 'bounded rationality' (Kahneman & Tversky, 1979) – so that it now looks a little like psychology. Nevertheless, the fundamentals of the argument just given are unaltered by these intellectual transfers. The economic method is still that approach which takes as its basic premise the thought that agents aim at ends, and that they do so in as rational a manner as is possible. That others have now adopted the economic method, or that economists impose certain caveats on their analyses, does not vitiate this essential feature of economics. Indeed, the fact that Gary Becker and George Akerlof, who brought the economic method to sociology, Kenneth Arrow, James Buchanan and Thomas Schelling, who undertook work foundational to modern political theory, and Daniel Kahneman, who introduced psychological ideas into economics, have all received Nobel prizes in economics is evidence of the fact that economics is the 'home' discipline of the analysis of rational – and 'nearly' rational - behaviour.

At this point, it's useful to specify the basic economic approach in the kinds of terms in which economists normally think. Necessarily, this involves a certain amount of formalism. We begin by supposing that

each person has a well-defined utility function, $U(\mathbf{x})$, which specifies the agent's preferences over the various outcomes he values. This is to say that the $U(\mathbf{x})$ function indicates how much utility the agent gets from each amount of the goods given in the vector, **x**. The agent also faces a set of constraints, $g(\mathbf{x}) \leq c$, which determines the range of options he has. The agent then attempts to solve the following problem. Choose **x** so as to:

$$\text{maximise } U(\mathbf{x})$$
$$\text{subject to: } g(\mathbf{x}) \leq c.$$

In words, the agent attempts to maximise his utility given the options he has available to him.

In the context of the interaction between two or more agents, economists deploy that particular methodological apparatus which is built for this specific purpose, viz., game theory. To see how that method works, let's suppose that there are two agents, A and B. A has a set of strategies available to him given by the vector, **a**, with $a_i \in \mathbf{a}$; and B has a set of strategies available to her given by the vector, **b**, with $b_i \in \mathbf{b}$. The utility of each player depends on the strategies adopted by both; i.e. we have:

$$U_a = U_a(a_i, b_i)$$
$$U_b = U_b(a_i, b_i).$$

A's problem is to choose $a_i \in \mathbf{a}$ so as to:

$$\text{maximise } U_a(a_i, b_i)$$
$$\text{subject to: } \sum p(b_i) = 1$$

where the constraint says that the beliefs that A has about how B will behave (which are given as subjective probabilities, p(.)) must be consistent (i.e. must add up to 1). In other words, A is supposed to choose the strategy that will maximise his utility given his beliefs about what B will do. B's situation is symmetrical.

Owing to a famous theorem by Nash, we know that it's possible for both players simultaneously to 'strategise and maximise' in a consistent manner. To see what this means, consider the following: let a_i' be a rational response by A to the play b_i' by B (i.e., given b_i', a_i' maximises the utility of A), and let b_i' by B be a rational response to the play a_i' by A (i.e., given a_i', b_i' maximises the utility of B), then (a_i', b_i') is a consistent set of strategies. In reference to Nash's proof of the fact that every game has at least one such set of consistent strategies, we call such a set a Nash equilibrium, and we usually denote it as: (a_i^*, b_i^*). If all players are rational, then we should expect only to observe Nash equilibrium outcomes, as any other outcome is inconsistent with both players acting rationally.

To see what this means, we can consider the following well known – hawk-dove – game. In this representation of the game, there are two agents: the USSR and the USA; and each agent has two strategies: Hawk (threaten war) and Dove (be peaceful); the payoffs that each nation receives are specified as follows: the payoffs on the left of each cell of the matrix are those

of the USSR, and the payoffs to the right are those of the USA:

		USA	
		Hawk	Dove
USSR	Hawk	0 0	1 5
	Dove	5 1	2 2

In this game, if both the USSR and the USA go to war (play Hawk), both get zero utility (since the world ends), if both pursue peace (play Dove), both get a utility of 2; however, if one threatens a 'hot' war and the other yields, the former 'wins' the Cold War and gets a payoff of 5, whilst the other loses and gets a payoff that's better than mutually assured destruction (i.e. 0) but is less than it would obtain under a mutual peace (i.e. 2) – i.e., it gets a payoff of 1. It's then a straightforward exercise to see that there are two Nash equilibria in this game: (Dove, Hawk) and (Hawk, Dove) (which are, respectively: the USSR yields to an American threat, and America yields to a Soviet threat). It's not consistent for both players to play Hawk or to play Dove since, if one player plays Hawk, the rational response of the other is to play Dove; and if one player plays Dove, the rational response of the other is to play Hawk. Hence, neither (Hawk, Hawk) nor (Dove, Dove) is a Nash equilibrium.

The prediction of the theory as to how the Cold War would play out is that one side would win (since these are the only possible Nash equilibria – i.e. the only

consistently rational outcomes). In fact, this is what was observed.

In the economic analysis of drugs usage in sports contests, we typically deploy a game theoretic approach to explain the conduct of the sportsmen for the obvious reason that sports contests are games in both the common and the theoretical meanings of that term. Before discussing the insights that game theory holds for the analysis of the issue at hand, however, I want to clear up one more point about the general character of the economic method.

The economic method just described is thought to have at least two possible uses. The first is as a description of the way actual agents behave; the other is as a set of norms that determine an ideal of rational conduct.

With regard to the first usage, there is plenty of criticism that economics is 'unrealistic' in the assumptions it makes and, therefore, in the conclusions that it reaches. There is a widespread perception (at least outside economics!) that this is generally the case; and there is also, by now, an extensive experimental literature which shows that people do not generally behave in a rational way. The literature on this matter was initiated by the work of Kahneman & Tversky (1979) amongst others, but it is now very diverse.

Since the empirical research has tended to contraindicate the rationality postulate, there has been a tendency to emphasise the alternative use of the

economic method as determining a sort of (ideal) logic of rational choice. The argument here is that, although economics can't very well explain what people actually do, it can explain what they would do were they rational – i.e., it can propose a kind of rational ideal to set against the reality of people's actual incompetence.

This bifurcation into descriptive – or 'positive' theory – and normative theory, however, is largely based on an epistemological error. All theory necessarily involves some degree of idealisation or abstraction, and all theories adopt some kind of idealised/abstracted entities that are reified in the theories' models (one thinks here, for example, of the role that the centre of mass plays in Newtonian mechanics). The question, then, is not whether the world strictly conforms to the picture generated by the models, but whether the models predict (or retrodict) changes in the world to a sufficient order of approximation given an initial state of play. Adapting this thought to the current context means that economics doesn't require people to be rational all of the time or even most of the time, it only requires that people be mostly rational most of the time; and that is probably of workable hypothesis.

II. The economic method and drugs in sport

In the context of sports competition, each agent is assumed to want to do the best he can, but he is constrained in this pursuit by his natural ability, the availability of time, a finite training budget and, of course, the behaviour of his competitors. Given those constraints, he works out a strategy that maximises his chances of winning. In what follows, we concentrate

solely on the constraints that the other players impose on a given athlete in his attempt to win. We do so as this is where the incentive to cheat lies in sporting competitions. Moreover, we shall assume that there are only two competitors as this allows us to focus on the fundamental methodological issues. The models can be generalised to many players.

There are a number of ways in which to model the behaviour of athletes-*qua*-agents even within the limits we have set ourselves. Perhaps the most intuitive model is given in the following game matrix:

		Athlete 2	
		don't cheat	cheat
Athlete 1	don't cheat	2 2	3 0
	cheat	0 3	1 1

In this representation of the situation in which the agents find themselves, both would prefer to compete in a drug-free – 'fair' – environment (and get utilities of 2) rather than in an environment in which both cheat (where both get utilities of 1); despite this common preference, however, both have an incentive to cheat. In fact, the incentives described in the given game are so strong that both will cheat *regardless of the actions of the other player*, and they will do so knowing that the end result will give each player a lower utility than he would have attained if both had chosen not to cheat.

To understand the motives for cheating in the above game, it's useful to consider the permutations of play a

little more carefully. In what follows, and for simplicity's sake, we use the first person pronoun to indicate one of the players. First, consider the case where the other player doesn't cheat, then I attain an advantage by cheating (since my odds of winning are increased), and my utility increases from 2 to 3. Secondly, if the other player cheats, then I have to cheat in response just to stay in the game. Specifically, if the other player cheats, I get no satisfaction from 'competing – and probably losing – with integrity' (my utility is 0 in that case); I'd rather cheat to give myself a chance of winning even though this win, should it occur, is tainted by my discreet drug use (the utility of playing and using drugs when all players are using drugs is 1). This situation is, of course, an application of the Prisoner's Dilemma to drug usage, and it is widely accepted in economics as a valid representation of the incentive to cheat in sports (see, e.g., Bird & Wagner, 1997).

In the Prisoner's Dilemma representation, each player has an incentive to cheat regardless of the conduct of the other player. Hence, cheating is a dominant strategy for each player and the unique Nash equilibrium involves both players cheating. It has, however, been argued that another representation better captures the nature of cheating (see Berentsen, 2002):

		Athlete 2	
		don't cheat	cheat
Athlete 1	don't cheat	2 2	1 0
	cheat	0 1	1 1

The argument for this payoff structure goes as follows. Both players would most prefer to compete fairly in a clean sport (and get a payoff of 2). If, however, the other player decides to cheat and I do not, I lose two units of utility relative to this optimum since, not only are my odds of winning reduced, I'm also now participating in a tainted sporting contest (hence my utility is 0). If, on the other hand, *I* cheat and the other player doesn't, my odds of winning are increased, but I regret the fact that I'm now participating in a tainted contest and that I'm the cause of the contest's being tainted (hence, my utility is also 1 in this case). Finally, if both players cheat, they lose a similar amount of utility relative to the mutual 'don't cheat' case since both are the cause of the sport's now being tainted and each has the same chance of winning as before (so each player has a utility of 1).

In this version of the game, each player wouldn't cheat if he were sure that the other wouldn't cheat – which reflects the fact that players have some degree of integrity (i.e. a greater degree of integrity than is conveyed in the Prisoner's Dilemma representation). However, each player cheats if the other is thought to cheat. Moreover, since cheating guarantees a certain payoff of 1 (as it assures the athlete of his competitiveness regardless of what the other player does), whereas not cheating carries the risk of getting 0 (if the other player cheats), there is an argument based on risk considerations for cheating (such arguments are called 'risk dominance' arguments).

The outcome in this representation of the incentive to cheat therefore seems similar to that given in the

Prisoner's Dilemma representation. However, mutual cheating, if it's observed in the above game, is premised on a much weaker set of incentives. For, in the above case, the athletes are at least open to the possibility of mutual honesty (i.e. mutually not cheating); which is to say that such an outcome is rational for both players in the sense that it is a possible Nash equilibrium. The difficulty lies in convincing the other party that each will not cheat. If such assurances are accepted, then cheating will not occur. Hence, if the above representation is a more accurate one than the Prisoner's Dilemma, we should expect to see less cheating as, at least in some cases, the agents carry out the 'don't cheat' Nash equilibrium. This game, for the historically interested, is an instance of Rousseau's Stag Hunt.

At this point, before drawing any general conclusions, I want to look at an intermediate case in which cheating is a dominant strategy for one player, and the other player is then motivated to cheat even though this is not a dominant strategy for him.

		Athlete 2	
		don't cheat	cheat
Athlete 1	don't cheat	2 2	1 0
	cheat	0 3	1 1

In this case, the payoffs for athlete 1 are as they were in the Prisoner's Dilemma, and for athlete 2, they are as they were in the Stag Hunt. Hence, athlete 1 will cheat regardless of what athlete 2 intends to do; and athlete 2, although he would be happy not to cheat if he

thought athlete 1 wasn't going to cheat, decides to cheat as well since he knows that athlete 1 will certainly do so. In this intermediate case, the payoffs in the game are not symmetrical, which represents the heterogeneous character of agents' preferences. And, since agents' ethical standards are incorporated in their expressed preferences, this representation captures agents' different ethical outlooks. Hence, this model shows that, while players may face different incentives to cheat, each may still end up doing so. In other words, the relatively unethical (athlete 1 types) are able to 'force' the relatively ethical (athlete 2 types) to behave just as unethically (i.e. to cheat). This is to say that relatively unethical players, by their intentional conduct, create incentives for relatively ethical players to cheat, and so drag down the whole game. In this kind of situation, although there may only be a few bad apples that are the cause of the problem, they rapidly infect the whole barrel. Their doing so, it may be said, sets off an 'epidemic of defection' (since it induces a wholesale switch from relatively ethical to relatively unethical behaviour). That being said, it needs to be emphasised here that the terminology used in this argument – of 'force', 'cause', 'infect', 'epidemic', etc. – is purely metaphorical; the real cause of the observed behaviour lies in agents' rational responses to the conduct of others playing the game, who themselves, of course, are also being rational.

The question now arises: what do these models show? First, they show that agents cheat because they have incentives to do so; which is to say that, given agents' own preferences and the expected conduct of their competitors, it is rational for players to cheat. This is

the fundamental insight into cheating offered by the economic method.

Second, the above models indicate that endemic cheating is a robust phenomenon; which is to say that cheating is a Nash equilibrium in a variety of games. Cheating will occur in Prisoner's Dilemma, Stag Hunt and hybrid games. In fact, it can be shown that cheating will occur in a yet wider variety of games when there are two or more agents. Hence, the economic method is able to model the real heterogeneity of attitudes to cheating, and the different interactions that thereby result. It also shows that cheating occurs even when some – and perhaps a plurality of – agents would be willing not to cheat, and would prefer not to cheat, if they thought others wouldn't; yet they do so because they fear that others will. Thus, the economic method accounts for both the observed prevalence of cheating and the widely expressed desire of sportsmen to eliminate it from sports. Such a phenomenon does not reflect a kind of schizophrenia or hypocrisy on the part of athletes, but is capable of being given a straightforward, rational explanation.

Third, the above models suggest, at least implicitly, how it might be possible to abate drug usage in sports. There are two avenues via which this might be done. The first is to reduce the rewards for cheating, and the second is to increase the rewards for not cheating. Generally speaking, the approaches currently adopted to reduce drug taking employ the former route. The use of regulatory bodies, negative lists of proscribed substances, fines, bans, and so on, exemplify this approach. The idea is to lower the payoff to

cheating by increasing the odds of being caught if one does so and by penalising convicted malfeasance. Given the arguments above, such an approach might be expected to have some effect on drug usage abatement.

Recently, however, arguments have been put forward that advocate a switch from reducing the payoffs to cheating to increasing the payoffs to not cheating. The argument for the suggested switch goes as follows. Suppose that a sport that is evidently drug-free benefits from increased customer demand and corporate sponsorship relative to sports that are evidently drug tainted. These pecuniary benefits might then be used to provide incentives for sportsmen to create an anti-drug culture in the sport in question. Specifically, the peak body of the sport might be able effectively to propagandise the 'virtues' of the sport's being drug free to the sport's participants by pointing out to them the costs of cheating and the financial benefits of their not doing so, both individually and collectively. It might then set up an educational structure along with internal investigative and reporting mechanisms to discourage drug usage. In such a fashion, it might be possible to rely on a mix of customer demand, media exposure, and the rational pursuit of self-interest on the part of the sport's peak bodies and participants to reduce drug usage.

Whether this latter method of drug use abatement is more likely to be effective than the alternative of relying on regulatory authorities to enforce proscriptions is an open question. The point here is not to investigate that issue but rather to point to the avenues for investigation which the economic method

naturally opens up. In particular, we note that the economic method asks: what incentives induce rational sportsmen to take performance-enhancing drugs in the first place? And what counter-incentives are most likely to be effective in abating their drug use? Or, equally, which regime of drug abatement is most likely to induce rational, self-interested individuals not to take drugs? This, it seems, is a not unreasonable way of approaching the issue of drugs in sports.

References and further readings

Aristotle (350BC [2004AD]) *The Nicomachean Ethics* (Adelaide: ebooks@Adelaide)

Bentham, J. (1825) *The Rationale of Reward* (London: J&HL Hunt)

Berentsen, A. (2002) "The economics of doping" *European Journal of Political Economy*, vol.18, pp109-127

Bird, E.J. & Wagner, G.G. (1997) "Sport as a common property resource" *Journal of Conflict Resolution*, vol.44 no.6, December, pp749-766

Ehrenberg, R.G. & Bognanno, M.L. (1990) "Do tournaments have incentive effects?" *Journal of Political Economy* vol.98 no.6, pp1307-1324

Haugen, K. (2004) "The performance-enhancing drug game" *Journal of Sports Economics* vol.5 no.1, pp67-86

Kahneman, D. & Tversky, A. (1979) "Prospect theory: an analysis of decision under risk" *Econometrica*, vol.7 iss.2, March, pp263-292

Kräkel, M. (2006) "Doping in contest-like situations" *mimeo*

Kräkel, M. & Sliwka, D. (2004) "Risk taking in asymmetric tournaments" *German Economic Review*, vol.5 no.1, pp103-116

Maennig, W. (2002) "On the economics of doping and corruption in international sports" *Journal of Sports Economics* vol.3 no.1, pp61-89

Milgrom, P. & Roberts, J. (1992) *Economics, Organization and Management* (Upper Saddle River: Prentice Hall)

Mill, J.S. (1863 [2004]) *Utilitarianism* (Adelaide: ebooks@Adelaide)

Nash, J. (1951) "Non-Cooperative Games" *The Annals of Mathematics*, vol.54 no.2, pp286-295

Preston, I. & Szymanski, S. (2003) "Cheating in contests" *Oxford Review of Economic Policy*, vol.19 no.4, pp612-624

Savulescu, J.; Foddy, B.; & Clayton, M. (2004) "Why we should allow performance enhancing drugs in sport" *British Journal of Sports Medicine*, vol.38, pp666-670

Symanski, S. (2003) "The economic design of sporting contests" *Journal of Economic Literature*, vol.41 no.4, December, pp1137-1187