Strategic Cost and ‘Matching Pennies’

Justin Smith

SFI WORKING PAPER: 1999-07-048

Abstract

Evidence supports the contention that humans find it costly to carry out some sorts of strategies. Such costs are unlikely to be observed directly, but various models have been proposed to represent them. We discuss a method for testing the empirical worth of these models.

Strategic cost and 'matching pennies'

Justin Smith
Nuffield College, Oxford, OX1 1NF, UK*

3 March 1999

In everyday life, people frequently select paths of action for their simplicity, as well as for the payoffs to be received if the paths are correctly followed.[1] For example, a stranger asking the way to some place might ask for simple directions, preferring an easily remembered route to a more complicated one which is slightly shorter. Or, as has been suggested by Radner [9], Axelrod [3] and others, players of repeated games may forego a slightly increased payoff in order to use a strategy which is simpler, and by implication cheaper.

Consider, for example, an iteration of a stage game with the following pure action payoffs,

2

		h	t
1	h	$1 \quad -1$	$-1 \quad 1$
	t	$-1 \quad 1$	$1 \quad -1$

. (0.1)

These correspond to payoffs of the game 'matching pennies', in which players simultaneously call either 'heads' or 'tails', one player winning if the calls match, the other winning if they differ. In the unique Nash equilibrium of the usual mixed extension of the game, each player uses each pure action with independent probability one half at each stage. If the game happens to be played this way, then each player is actually indifferent between all possible sequences of actions. Furthermore, neither player can improve her expected payoff by 'second guessing' or attempting to predict what moves her opponent will play. However, there is considerable evidence from a number of experimental studies that second guessing human opponents is typically profitable, suggesting that human players of matching pennies do not in fact act so randomly. For instance, Hagelbarger [6] built a small robot known as SEER (SEquence Extraction Robot) which played the game against a variety of opponents. SEER was designed to recognise whether recent play by its opponent conformed to one of the periodic sequences

$$hhhhhh...$$

*Correspondence to justin.smith@economics.ox.ac.uk

[1] I would like to thank John Miller, Scott Page and participants in the 1998 Santa Fe Graduate Workshop in Computational Economics for their help and suggestions.

$$tttttt...$$
$$hththt...$$
$$hhtthh...$$

If SEER had won at least half the stages already played during a game, it would play a best response based on the assumption that any sequence would be continued. Otherwise it would play randomly. Even this simple device performed significantly better than a completely random player, against a significant majority of its opponents, despite the majority of its opponents being mathematicians, computer scientists or information theorists, who were well aware of the formal properties of the game described above.[2] One interpretation of these results is that the human players failed to work out their Nash strategy, but given their skills this seems implausible. Moreover, the cost of working out the Nash strategy should not increase with the number of repetitions of the game, at least if the relevant stage game is just as illustrated by 0.1, whilst the losses that may accrue from the use of some other strategy will typically increase with the number of repetitions of the stage game. Hence so long as the Nash strategy is not infinitely costly to compute, 'sensible' human players should compute it for use at some point during sufficiently long iterations of the game. Another interpretation is that, knowing that they were playing against a robot which was not necessarily playing the Nash strategy, they tried to second guess the robot, and failed. This explanation too seems implausible by itself, since the players could at least have switched to playing the Nash strategy, once it became evident that they were failing to second-guess the robot. A more plausible interpretation of Hagelbarger's results would seem to be that the humans found some strategies to be more costly to implement than others, and that in order to avoid the more costly ones to some degree, human players were prepared to use somewhat predictable strategies, and accept some losses.

If the latter explanation has a basis in fact, then games like matching pennies may provide a convincing illustration of the benefits of explicitly modelling the costs of carrying out strategies in games. These costs, which we refer to as strategic costs, can be modelled simply by formulating a game which is otherwise identical, but has a different payoff function. Thus, if a game is described to have a set of strategies A and payoff function $\Pi(A)$, strategic costs are assumed to be given by a function $C(A)$ such that the players of the game actually consider themselves to be playing a game with payoff function $\Pi(A) - C(A)$.

The case for being sure to include strategic costs in payoff functions has obvious similarities to the imperative to pay close attention to specification problems in applied econometric analysis. Indeed, the former may be seen as one instance of the latter. However, whilst it is standard for analysis of specification problems in econometrics to be carried out with close reference to the data being modelled, the literature on strategic cost has typically paid scant attention to its raison d'être, the failure of some 'canonical' payoff matrices to 'explain' the outcomes that are typically observed in the games they model. One result is that few attempts have been made to compare

[2] The present author built a slightly more sophisticated version of SEER at the 1998 graduate workshop in computational economics at Santa Fe, with similar results.

empirically the worth of the various systematic representations of strategic cost that have been developed. Thus, whilst we have many alternative systems of this kind, including though by no means limited to those proposed in [1], [2], [4], [7], [8], [10], and [11], we have no well-developed empirical critique of these systems.

This makes the game of matching pennies particularly interesting, since, if a player does play the game so as to minimise a sum of losses and strategic costs, then the better her strategic costs are understood by her opponent, the higher, one might expect, the opponent's payoff should be. Thus, by pitting a variety of automated players, each one programmed to behave as though it believes its opponents to face strategic costs as implied by some particular system, against a variety of human opponents, it may be possible to fit models of strategic cost and to achieve an empirical critique of some of the various systems. Perhaps also, by pitting human players against one another, we could assess their relative strategic costs.

If our interpretation of Hagelbarger's results is correct, and if we can show that it is possible to fit models of strategic cost, then the practical applications of the method could be many and widespread. Some of these are indirect applications. The development of theories of strategic cost could obviously benefit. So too could those parts of evolutionary game theory that rely on some notion of strategic cost or complexity to explain the relative likelihood of different sorts of perturbation in strategy choice. This applies both to the analysis of the evolutionary stability of strategies, in which it is usual to assume that the likelihood of a perturbation that results in a player using any particular strategy is a function of the complexity in some sense (and by implication the perceived cost) of that strategy, and also to more experimentally directed work, in which the stage game strategies of players in a repeated game are allowed to evolve on a set of candidate strategies, with the likelihood of evolution to any particular strategy being a function of the complexity of that strategy. There may also be many more direct applications, to situations in which, as in matching pennies, players can benefit at one another's expense if they correctly predict one another's actions. For example, when policy-makers use their predictions about aggregate demand and output to guide stabilisation policy, they succeed at the expense of workers or firms, whose real revenues turn out to be smaller than expected. Workers and firms have clear incentives to prevent this, which they may be able to do if they correctly anticipate policymakers' predictions. Many macroeconomists argue that given the incentives of firms and workers, policymakers can not confound the predictions of firms and workers to a greater extent than their own predictions are confounded. But the situation has obvious similarities to a repeated game of matching pennies, and if in such a game it is possible for a more sophisticated player, perhaps one who faces lower strategic costs, to win consistently against a less sophisticated player, then we can construct an argument for (and perhaps some specific guidelines towards) the use of stabilisation policy. There may be similar applications to the formulation of policy based on microsimulation. For example, it is common practice for tax inspectors to concoct likely profiles of tax dodgers to help them pick out individual returns for audit. Taxpayers (at least those ones filing false returns) would like to ensure that their own returns do not match these profiles. The situation is rather like a repeated game of matching pennies, in which one player uses some predictive algorithm to guess the move that the other player will make at each stage. If there are no strategic costs then in the unique Nash equilibrium the inspectors would not use any predictive algorithm and would select returns for audit completely at random. The efficacy of profiling may depend on the taxpayers facing strategic costs, and the use of profiling may provide empirical support, as well as a potential application, for models of strategic cost. Similar applications may be found in a variety of other situations, such as the selective monitoring of effort in a principal-agent relationship, that are not so commonly used as applications for microsimulation. Yet other applications, and indirect empirical support, may be found in the arena of competitive sports. For example, in baseball the pitcher must decide, with each throw, which of several pitches to use. The batter must guess which pitch the pitcher will use on each occasion and, in general, the payoff of the batter is higher, and the payoff of the pitcher is lower, when the batter guesses correctly. By counting the number of pitches available to the pitcher, it is possible to estimate the proportion of pitches that the batter should expect to anticipate correctly, if neither player faces any strategic costs, but the actual picture appears to be more complicated, and how many players, pitchers or batters, would agree that they use the implicit Nash strategies of complete randomisation? It is easy to find similar situations in other sports that are in practice commonly seen as tests of a player's mental aptitude and ability to outwit other players, but which, without something like an analysis of strategic costs, might be modelled be modelled by a game that allows no representation of such skill. For example, the question of where to direct a shot in a game of tennis or badminton, the question of which direction to send a penalty in soccer, or the question of whether to overbid in a game of bridge, or bluff in a game of poker, can all be seen in this light.

Can an experimental analysis of the size and any systematic structure of players' strategic costs be carried out by analysing play of games like matching pennies? If so, can the results be applied to improve players' choice of strategies in such games? Our aims in this paper are quite limited, but can be seen as a first step towards answering these questions. In the first three sections of the paper, we discuss some obstacles to the analysis of strategic costs in penny-matching games, and suggest ways to overcome or circumvent these obstacles. In section four we analyse the results of some simple experimental games played between humans and certain machine players. Section five concludes.

1. Costs of randomisation

The case for reformulation of payoffs to include strategic costs has most often been made with regard to pure strategies, but it is quite plausible in the light of research regarding people's choices under uncertainty that players incur particular costs when carrying out non-degenerate mixed strategies that are not accounted for by the payoff function of the standard mixed extension of a game. However, modelling the costs of mixing may change payoffs in such a way as to introduce particular difficulties to the analysis of games. In particular, a game whose payoff function incorporates costs of mixing may have Nash equilibria only in pure strategies, and may not have any Nash equilibria at all. On the other hand, if such a game has Nash equilibria other than in pure strategies, then the game may not preserve all or even any of the equilibria of the game without strategic costs.

Suppose

$$\Gamma = \langle A_1, ..., A_N, \Pi_1, ..., \Pi_N \rangle$$

is a game with N players, a set of actions A_i for each player i, and a function $\Pi_i : A_1 \times ... \times A_N \to \mathcal{R}$ for each player i that gives the payoff consequent to i from each profile of the players' actions. Allow that the set of actions available to each player i also includes all her mixed strategies, or probability distributions over the elements of A_i. Assume further, if p_j is a probability distribution defined over the elements of A_j, that

$$\Pi_i (p_1, ..., p_N) \equiv \sum_{(a_1,...,a_N) \in (A_1 \times ... \times A_N)} p_1 (a_1) ... p_N (a_N) \Pi_i (a_1, ..., a_N) \ .$$

In other words, the payoff due to use of a mixed strategy profile is, as usual, the expected payoff consequent on that profile. Let $\Delta (A_i)$ denote the set of probability distributions defined over the elements of A_i. The profile $(p_1, ..., p_N)$ is called a Nash equilibrium of Γ iff the mixed strategy p_i maximises the payoff of player i, with respect to the elements of $\Delta (A_i)$, for each player i. Let p_j be called a non-degenerate probability distribution over A_j iff

$$p_j (a_j) < 1$$

for all $a_j \in A_j$.

Now, suppose

$$\Gamma^* = \langle A_1, ..., A_N, \Pi_1, ..., \Pi_N, c_1, ..., c_N \rangle$$

is a game defined as Γ, except that for each player i, $c_i : \Delta (A_i) \to \mathcal{R}$ is a function giving the cost to i of using any of her available actions, and the payoff due to i from use of a profile $(p_1, ..., p_N)$ is defined to be

$$\Pi_i (p_1, ..., p_N) - c_i (p_i) \ .$$

Call $c_i (.)$ the strategic cost function for player i in Γ^*. Say that randomisation is costly for j iff j is a player such that $c_j (p_j) > \sum_{a_j \in A_j} p_j (a_j) c_j (a_j)$ for all non-degenerate probability distributions $p_j \in \Delta (A_j)$. Say that the costs of randomisation are zero for j iff j is a player such that $c_j (p_j) = \sum_{a_j \in A_j} p_j (a_j) c_j (a_j)$ for all probability distributions $p_j \in \Delta (A_j)$. Then,

Proposition 1.1. *If p_i' is a non-degenerate probability distribution over A_i in Γ^* and randomisation is costly for i, then p_i' is not a best response to any profile of mixed strategies in Γ^*.*

Proof. Suppose the contrary. Then

$$p_i' \in \arg \max_{\phi_i \in \Delta (A_i)} \{\Pi_i (p_1, ..., p_{i-1}, \phi_i, p_{i+1}, ..., p_N) - c_i (\phi_i)\} \qquad (1.1)$$

for some profile $(p_1, ..., p_{i-1}, p_{i+1}, ..., p_N)$. But

$$\Pi_i (p_1, ..., p_{i-1}, p'_i, p_{i+1}, ..., p_N) = \sum_{a_i \in A_i} p'_i (a_i) \Pi_i (p_1, ..., p_{i-1}, a_i, p_{i+1}, ..., p_N)$$

and $\sum_{a_i \in A_i} p'_i (a_i) = 1$, so for any $a_i^* \in A_i$ such that $p'_i (a_i^*) > 0$,

$$\begin{aligned} \Pi_i (p_1, ..., p_{i-1}, p'_i, p_{i+1}, ..., p_N) &= \frac{1}{p'_i (a_i^*)} p'_i (a_i^*) \Pi_i (p_1, ..., p_{i-1}, a_i^*, p_{i+1}, ..., p_N) \\ &= \Pi_i (p_1, ..., p_{i-1}, a_i^*, p_{i+1}, ..., p_N) \ . \end{aligned}$$

Since randomisation is costly for i, $c_i (p'_i) > \sum_{a_i \in A_i} p'_i (a_i) c_i (a_i)$. Letting

$$a_i^* \in \arg \min_{a_i \in A_i} \{c_i (a_i) \mid p'_i (a_i) > 0\} \ ,$$

it is then clear that

$$c_i (p'_i) > \frac{1}{p'_i (a_i^*)} p'_i (a_i^*) c_i (a_i^*) = c_i (a_i^*) \ ,$$

and thus that

$$\Pi_i (p_1, ..., p_{i-1}, a_i^*, p_{i+1}, ..., p_N) - c_i (a_i^*) > \Pi_i (p_1, ..., p_{i-1}, p'_i, p_{i+1}, ..., p_N) - c_i (p'_i) \ .$$

It follows that

$$p'_i \notin \arg \max_{\phi_i \in \Delta(A_i)} \{\Pi_i (p_1, ..., p_{i-1}, \phi_i, p_{i+1}, ..., p_N) - c_i (\phi_i)\} \ ,$$

which contradicts 1.1. ■

Corollary 1.2. *If p'_i is a non-degenerate probability distribution over A_i in Γ^* and randomisation is costly for i, then any profile $(p_1, ..., p_{i-1}, p'_i, p_{i+1}, ..., p_N)$ is not a Nash equilibrium of Γ^*.*

Corollary 1.3. *If Γ_T^* is a T-times repetition of Γ^*, if S_i is the set of i's pure strategies in Γ_T^*, if randomisation is costly for i, and if p'_i is a non-degenerate probability distribution over S_i, then any profile of mixed strategies $(p_1, ..., p_{i-1}, p'_i, p_{i+1}, ..., p_N)$ is not a Nash equilibrium of Γ_T^*.*

It follows straightforwardly that if randomisation is costly for the players of a game, the game may not have an odd number of Nash equilibria, even if it has a finite number of pure strategies, and that such a game may not have any Nash equilibria at all. If the costs of carrying out pure strategies in such a game are not always zero, then depending on the game and the costs there may be more, fewer or the same number of equilibria as there would be if all strategies could be carried out costlessly. To illustrate each case, consider a game with payoffs

1		l		r	
	u	4	-1	2	1
	d	-1	1	1	-1

,

in which randomisation is costly for both players. If each pure strategy can be carried out at zero cost then the game has a unique Nash equilibrium, (u, r). If we then introduce $c(u) = 3$ then the game has no equilibria. If, alternatively, we introduce $c(r) = 2$, then there are two equilibria; (u, l) and (u, r).

These possibilities, particularly the nonexistence of Nash equilibrium, present obvious obstacles. However, thinking about the conditions under which the game may be played calls attention to a curious requirement for an action profile to be a Nash equilibrium of a game. Suppose a game which consists of a single play of matching pennies, with pure strategy payoffs as represented by 0.1, in which randomisation is costly for each player. Even if a player, say, actually chooses to play h with probability one, is it reasonable to suppose, as the concept of Nash equilibrium would require, both that the other player is, necessarily, fully aware of this choice and that this awareness is common knowledge? We believe not, but what is it reasonable to suppose instead that each player believes about what the other player is playing?[3]

One possible condition is that players hold prior probability distributions on one another's action choices that happen to be the same as the distributions implied by a Nash equilibrium of the version of the game in which costs of randomisation are zero. If we denote the beliefs of a player i about the action chosen (or to be chosen) by player j by the probability distribution p_{ij}, then this condition can be restated as

$$p_{ij} \in \arg \max_{\phi_j \in \Delta(A_j)} \left\{ \Pi_j \left(\phi_j, \phi_i\right) - \sum_{a_j \in A_j} \phi_j\left(a_j\right) c_j\left(a_j\right) \right\}, \tag{1.2}$$

for all i, j, $i \neq j$, where the profile of distributions $\phi_1, ..., \phi_n$ constitutes a Nash equilibrium of the version of the game in which costs of randomisation are zero. Clearly there always will be such a profile of distributions, since the variant of the game in which randomisation is costless always has some Nash equilibrium. Moreover, in such an equilibrium each player for whom randomisation is costly is actually indifferent between the use of all (pure or mixed) strategies in the support of her mixed strategy. So long as the distributions $\phi_1, ..., \phi_n$ are common knowledge, then in the version of the game in which randomisation is costly, each player would still be indifferent be-

[3]Interestingly, Von Neumann & Morgenstern [12] apparently thought differently to us. With reference to the game of matching pennies, they wrote (section 17.2.1),

> "one important consideration for a player in such a game is to protect himself against having his intentions found out by his opponent. Playing several different strategies at random, so that only their probabilities are determined, is a very effective way to achieve a degree of such protection: By this device the opponent cannot possibly find out what the player's strategy is going to be, since the player does not know it himself. Ignorance is obviously a very good safeguard against disclosing information directly or indirectly."

Indeed it is, if randomisation is not costly. So far as we are aware, Von Neumann & Morgenstern did not explicitly consider the possibility that this might not be so.

tween the use of each pure strategy in the support of her mixed strategy, even though, of course, she would not use the mixed strategy itself.

This condition is not the only one that might be imposed to answer the question about players' beliefs, and it may not always provide the basis for a plausible answer to such questions. Apart from expediency, the main argument we offer for it is in general only suggestive. That is; even in a game like matching pennies where the only Nash equilibria are in non-degenerate mixed strategies, the incentive to randomise stems only from a desire to obscure the actual actions to be played until after the other player has played. If players are not implicitly aware of one another's strategy choices, and if each pure strategy in the game can be used at zero cost, then the fact that a player is in fact using a pure strategy rather than randomising could be kept hidden from opponents, because any observed action sequence that results from the use of some non-degenerate mixed strategy could also have been produced by some pure strategy. Each player could use some pure strategy without fear of his opponent becoming aware of that strategy, even in an arbitrarily long repetition of the game.

On the other hand, there are several obvious objections to the imposition of condition 1.2. Firstly, if randomisation is not costly for all players of a game then some player may have a strong incentive to randomise, and to use strategies that he would certainly not use in the version of the game in which randomisation is costless. For example, consider a finite T-times iteration of the prisoners' dilemma game;

2

1		c		d	
	c	3	3	0	4
	d	4	0	1	1

If the players of this game cooperate by playing trigger strategies, with the potential for immediate punishment, then neither player has an incentive to deviate until the final stage, and if 1 can only achieve such deviation at a cost of more than one unit, he will not choose to do so. If it is known that 1 will not deviate, then of course neither player has an incentive to deviate until the last stage, and cooperation until then may be possible in equilibrium. However, whilst it is difficult to be absolutely sure of deviating at the last stage, it might be quite easy to do so with a little less reliability. Suppose, for example, that although 1's pure strategy 'use a trigger strategy until the final stage, then play d' costs two units, 1's mixed strategy 'use a trigger strategy until the final stage, then with probability 0.9 play d' only costs half a unit. Then, clearly, 1 would have a strong incentive to use the latter strategy in place of either the former one or the straightforward 'grim' trigger strategy, and knowing this, it may be that neither player would ever play c in equilibrium during the game. However, the force of this objection is mitigated in the case of a game played between a machine and a human, for we can arrange the play of the machine, though not that of the human, to be consistent with optimal play by a player who faces no costs of randomisation. The possibility that a human player may prefer to play strategies with a 'small' amount of randomisation, rather than pure strategies, would remain.

Secondly, even if randomisation is costly for both players of a game, why should 1.2 be common knowledge? Since our reasoning began by questioning why players' actions should be common knowledge, it is natural to go further and wonder why their beliefs about one another's actions should be common knowledge, or, for that matter, their beliefs about one another's beliefs about one another's actions, and so forth. But if, taking this to its logical conclusion, we conclude that in general none of these beliefs are common knowledge, then it is hard to see how a stronger equilibrium condition than rationalisabilty could be used to answer the original question about what beliefs are plausible. Unhelpfully, in a game of matching pennies, even if some pure strategies are costly, the set of rationalisable equilibria may be large and may even include all pure strategy profiles. Thus it would not provide a very useful foundation for an analysis of the game. However, the force of this objection too may be mitigated in the case of a game played between a machine and a human. For, the machine can easily be programmed to play a strategy that would be optimal if it believed that 1.2 was common knowledge. If the human player is convinced that the computer is indeed playing in this way, then it would become optimal for the human player too to play as if 1.2 was common knowledge. Unfortunately, in the case of a game between two human players, the objection may be harder to dispose of. And if the objection turns out to be correct then players' action choices are likely to depend on attempts to outguess one another, which may involve complex reasoning about players' higher order beliefs. If this is so, and one player performs badly relative to another, it will require careful experimental design to be able to decide whether the failing player faced higher strategic costs or was less able to carry out higher order reasoning tasks.

Thus, the assumption that 1.2 is common knowledge but that players' chosen actions are not - or equivalently the assumption that randomisation is costless in the sense defined above - is a convenient and perhaps not too problematic device to ensure the existence of equilibrium. Moreover, it simplifies considerably the task of analysing gameplay, since it allows us effectively to ignore any costs of randomisation that players face.

2. Costs of pure strategies

Suppose that, including strategic costs, the pure action payoffs of a matching pennies game are as represented by the following table;

2

1		h		t	
	h	$1-c(h)$	-1	$-1-c(h)$	1
	t	$-1-c(t)$	1	$1-c(t)$	-1

, (2.1)

where $c(h), c(t) \in \mathcal{R}$, and randomisation is costless. To compute equilibrium points of the one-shot game is quite easy. If we let p_i describe the probability with which each player chooses action h, then we can find each player's Nash strategies (there will generally only be one for each player) by maximising the expected payoffs,

$$-p_2 p_1 - (1-p_2)(1-p_1) + p_2(1-p_1) + (1-p_2)p_1$$

and

$$p_2 p_1 (1 - c(h)) + (1 - p_2)(1 - p_1)(1 - c(t))$$
$$-p_2 (1 - p_1)(1 + c(t)) - (1 - p_2) p_1 (1 + c(h)) \ ,$$

with respect to p_2 and p_1 respectively, subject to the constraints $0 \leq p_1 \leq 1$ and $0 \leq p_2 \leq 1$. Doing so yields

$$p_1 = \begin{array}{ll} \frac{1}{2} & \text{if } 0 < 2 + c(h) - c(t) \leq 4 \\ 1 & \text{if } 2 + c(h) - c(t) \leq 0 \\ 0 & \text{if } 2 + c(h) - c(t) > 4 \end{array} \tag{2.2}$$

and

$$p_2 = \begin{array}{ll} \frac{2+c(h)-c(t)}{4} & \text{if } 0 < 2 + c(h) - c(t) \leq 4 \\ 0 & \text{if } 2 + c(h) - c(t) \leq 0 \\ 1 & \text{if } 2 + c(h) - c(t) > 4 \end{array} \ . \tag{2.3}$$

Heuristically, unless player 1 has sufficiently asymmetric costs as to have a dominant strategy, 2 will choose a strategy such that 1's best response is to randomise completely. This implies that unless 1 has sufficiently large costs as to have a dominated action, 2 can not in any Nash equilibrium earn a higher payoff than would be achievable if 1 had no strategic costs. If play is actually in equilibrium, 2 is more likely to 'win' only if 1 has a dominated strategy. However, in this case, as in a prisoners' dilemma game, there are outcomes that Pareto dominate the equilibrium outcome. For, the equilibrium outcome may involve 1 playing her more costly action with strictly positive probability, in which case both players could be made better off by having 1 play her more costly action with lower probability, and having 2 play with slightly greater probability the action that 'wins' against this action of 1. For example, if 1 faces costs of $c(h) = 1$ and $c(t) = 0$, and randomisation is costless, then the unique Nash equilibrium of the game is $p_1 = \frac{1}{2}$ and $p_2 = \frac{3}{4}$, but this outcome is Pareto dominated by all the points in the shaded area in figure **??**, which maps the possible action combinations in the game.[4] One might expect, as in a prisoners' dilemma game, that although there is a unique Nash equilibrium, players will in experiments somehow focus on one of the Pareto dominant action profiles. The Pareto dominant profiles actually comprise the right-hand border of the shaded triangle in figure **??**, in other words 2 plays t with zero probability in any Pareto dominant action profile. What outcome would one, as an experimentalist, expect to see? On one hand, there is a unique Nash equilibrium, but on the other hand there are a range of profiles that Pareto dominate the equilibrium. One might expect players to focus on one of the Pareto dominant profiles. But note that to do so so requires a high degree of trust on the part of 2, who must believe that 1 will not simply play $p_1 = 0$, despite knowing that it would be in 1's selfish interests to do so, and that it may be difficult even after the fact to establish whether 1 did in fact deviate. Moreover, the example assumes that randomisation is costless. If in fact randomisation is costly, then 1 could have an even stronger incentive to deviate to play of $p_1 = 0$; added to which we could hardly

[4]The figures are not included in this printing.

rely on the device, suggested in the previous section, of assuming that 1 actually uses a pure strategy but that it is not common knowledge. Of course, 2 can remove any part of the selfish incentive that 1 has to deviate from an outcome that Pareto dominates the Nash equilibrium; by setting $p_2 > 0$ and moving the outcome towards the Nash equilibrium profile; and a third sort of possible prediction is a profile that is somewhere between a Pareto dominant profile and the Nash profile. The question of what would be a sensible prediction in a game with such a pattern of strategic costs is a difficult and interesting one that could itself be investigated by experimental means. But an answer to this question is required in order to test any hypothesis about the systematic structure of strategic costs, and although a joint hypothesis test is quite feasible, this places an extra burden of explanation on the data. The situation is likely to be further complicated by the additional need to estimate the size of the strategic costs, which are unlikely in practice to be known a priori.

Finding equilibria of the repeated game could be much harder. Suppose that the game is repeated, but played under similar conditions, with costless randomisation and only player 1 having costs of carrying out strategies. Let Γ_T denote a T-times repetition of the matching pennies stage game with pure action payoffs (exclusive of strategic costs) represented by 0.1. One important problem to overcome is how to include strategic costs in the representation of the repeated game. In the case of the one-shot game, we could simply subtract them from the appropriate final payoffs, and one option in the case of the repeated game is to represent the game in strategic form and, similarly, subtract strategic costs from the appropriate final payoffs. However, it is neither necessary nor obvious that players facing strategic costs do in fact pay a once-for-all charge for the strategies that they use. Rubinstein [10] argues that in at least some circumstances it makes more sense to speak of a flow of costs, which could be interpreted, for example, as rental payments for the use of a machine capable of carrying out the strategy. Even if players do pay a once-for-all charge, it is not immediately clear whether it would be more appropriate to speak of it being incurred at the outset of the game or at the end. These considerations are irrelevant to the equilibria of the game in strategic form. However, the costs of playing particular actions at any particular stage might well depend on the play path taken before arriving at this stage. For example, it may be less costly for 1 to play h if he played h at the previous stage than if he played t at the previous stage. If this is so, and if in addition players can affect the amount of costs they will ultimately pay by changing their strategy choices during play of the game, then the manner in which strategic costs are incurred might nonetheless be significant to the outcome of the game, even if we assume that one player faces no strategic costs. The key to this is that if we admit that some player may face strategic costs, then unless special restrictions are placed on those costs, we can not assume that the game is strictly competitive, and if it is not strictly competitive then the game may share some rather awkward characteristics of other games, particularly if the game turns out to possess multiple Nash equilibria. To illustrate, suppose that there exists a Nash equilibrium of the repeated game, in which player 2 forces 1 on to a play path on which 2 will ultimately win more stages.[5] It may be the case that the establishment or maintenance of such

[5]It may even be possible for player 1 to force player 2 on to a play path on which 1 will ultimately

an equilibrium would require 2 to hold beliefs about the out of equilibrium actions of 1 that are not 'credible' - for example, the equilibrium may not be subgame perfect or sequential - but the question of credibility can not be addressed if the game is represented in strategic form. Even if the game is represented in extensive form, the job of deciding exactly what beliefs are 'credible' is subject to its own well-known difficulties. Furthermore, if the game is represented in extensive form, the manner in which players are charged for their use of strategies (as well as the actual costs) may affect which possible beliefs are 'credible'. For example, if players pay a once-for-all charge at the outset of the game for the strategies they use, then the payment of such a charge might constitute an effective precommitment to the use of a particular strategy. Such precommitment might not be possible if players are charged at the end of the game, on the basis of the strategies they have actually used, or if they are charged a small amount at each stage, on the basis of the strategy they are using at that point, if they are thereby allowed to change their strategy choices as the game progresses. A specific instance of this sort of contrast exists between the model of Rubinstein [10] and that of Abreu & Rubinstein [1].

It may be, of course, that none of these possible problems; not the existence of 'non-credible' equilibria in a repeated matching pennies game with strategic costs, nor opportunity for precommitment, nor even the game not being strictly competitive; is particularly likely to occur in any actual version of the game that we choose to study. In particular, none of these difficulties is likely to occur if strategic costs are sufficiently small relative to the underling payoffs of the game, and in principle we can engineer the strategic costs to be relatively small, by increasing the underlying payoffs of the game. If such difficulties are judged sufficiently unlikely to occur, then it may be plausible to ignore them. Computation of optimal strategies may then be a much simpler task. However, we have not computed such strategies, indeed we have not described very precisely the difficulties that might be assumed away! For what it's worth, our conjecture about the Nash strategies of a version of the game in which strategic costs are sufficiently small, is stated here as Conjecture 2.1. This conjecture is based on the Nash strategies of the one-shot game, but nonetheless allows that player 1's strategic costs in any stage game depend on the path of play taken to reach that stage game. Let Γ_T^* be a T-times repetition of Γ^*, let σ_i be a strategy for player i in Γ_T^*, let and let s_τ denote the history of play that has occurred up to and including stage $\tau - 1$ of Γ_T^*. Let $c\,(x \mid s_\tau)$ be the utility cost to player 1 of playing action x at stage τ after history h_τ has occurred, any cost incurred being subtracted from 1's final payoff.

Conjecture 2.1. *Any Nash equilibrium of* Γ_T^* *must require play of the following*

win more stages, even if 2 faces no strategic costs, if 1 carries out some strategy that was accorded zero prior probability by 2. For example, SEER could be interpreted as the Nash strategy choice of player 2 if 2 believes with probability one that the strategy chosen by player 1 must be drawn from a certain rather limited set. However, there are strategies with which player 1 could beat SEER, although player 2 would accord zero probability to the evnt of one of them being used.

actions at each stage;

$$p_1 = \begin{array}{ll} \frac{1}{2} & \text{if } 0 < 2 + c(h \mid s_\tau) - c(t \mid s_\tau) \leq 4 \\ 1 & \text{if } 2 + c(h \mid s_\tau) - c(t \mid s_\tau) \leq 0 \\ 0 & \text{if } 2 + c(h \mid s_\tau) - c(t \mid s_\tau) > 4 \end{array}$$

and

$$p_2 = \begin{array}{ll} \frac{2+c(h \mid s_\tau)-c(t \mid s_\tau)}{4} & \text{if } 0 < 2 + c(h \mid s_\tau) - c(t \mid s_\tau) \leq 4 \\ 0 & \text{if } 2 + c(h \mid s_\tau) - c(t \mid s_\tau) \leq 0 \\ 1 & \text{if } 2 + c(h \mid s_\tau) - c(t \mid s_\tau) > 4 \end{array} .$$

Given this conjecture, the computation of optimal strategies at each stage of the game is a relatively simple business, at least once appropriate values for $c(h \mid s_\tau)$ and $c(t \mid s_\tau)$ have been computed. Of course, as in the one-shot game, it might be the case that a Nash equilibrium profile would not be a good prediction of the strategies chosen by actual players, and whilst choosing a Nash action to play against an opponent who expects something else will not lead to a lower payoff, it may do in the repeated game, if the opponent revises her expectation and an opportunity to play action profiles that Pareto dominate Nash profiles in the continuation game is thereby lost.

On the other hand, unless analysis must yield either one equilibrium or several equilibria with similar payoffs, and unless play conforms with an equilibrium profile, any inference that could be drawn from observed play, concerning the strategic costs faced by some player, may have to depend also on a theory of equilibrium selection. Since a comprehensive theory of equilibrium selection is currently beyond our understanding, any such inference would be insecure.

One way to avoid some of the more difficult aspects of equilibrium selection problems, particularly those that depend on the dynamic effects of interaction between the players, can be achieved in practice by ensuring that the game is played without any ongoing interaction between the players, whereupon even the repeated game can be modelled in strategic form. The possibility of interaction can easily be removed from actual play by ensuring that each player does not observe any of the other player's actions until after the end of the game and the assignment of final payoffs. In fact, in the case of a game played between a computer player of our own design and a human player whose strategic costs are under study, only a rather weaker condition is required. We can set the computer player always to play a Nash strategy, and allow it to observe the actions played by its opponent, and merely prevent the human player from observing the actions played by the computer player. This approach very neatly avoids some of the thornier equilibrium selection problems, but it does so by imposing rather drastic limitations on the manner in which the game to be studied must be played, and on the set of strategies available to the player whose strategic costs are the object of the study. In a T-times repeated game of matching pennies, in which a player can observe all the actions played by herself and her opponent, that player has 2^{2^T} pure strategies. By contrast, if in the same game a player can only observe her own past moves, then she only has 2^T pure strategies, or $\frac{1}{2^T}$ times as many. This is still a very large number for even moderately-sized T, larger, perhaps, than the number of pure strategies between which the player might actually choose. But many strategies that are commonly thought of as being archetypically simple or easy to play do not fall in this category, for example; tit-for-tat and trigger strategies. Restricting a player to use only strategies that do not condition on some other player's actions is not trivial. Nonetheless, we judge that the greater tractability of this special case makes it the most suitable one for study at present.

Besides, the restriction leaves open to study some instances of games that resemble in certain respects the general case in which players can react to one another's actions at each stage. For example, to analyse the case of a repeated game of matching pennies played between a computer and a human player, the restriction could require that the human player could not observe the actions played by the computer, at least until after the end of the game. But we could analyse, consistent with the restriction, a game in which there are three players; a computer and a human who play matching pennies against one another, and another human player, whose actions are irrelevant to final payoffs but are observed by both other players, where the play is arranged so that the two human see each others moves at each stage and believe themselves to be playing matching pennies against each other. Whilst this is not quite the same as allowing the human whose strategic costs are being studied to observe the actions of the player that it is 'really' playing against, it does at least allow that human to choose from a similar set of strategies.

Also, there are many real world games, which might provide useful applications, that might be represented by the special case in which one player has limited ability to observe moves made by the other player, and also faces much larger strategic costs. For example, consider the game described by the following passage.

> Decades of seizures [of contraband from road hauliers] have bequeathed [to customs inspectors] an encyclopaedia of clues, which are fed into a computer that churns out continuously updated and refined profiles of likely smugglers.
>
> No detail is too small. A truck with a crooked exhaust? Evidence of neglect; maybe the driver was too busy picking up an illegal load to fix it. Check him. An out of date tax disc? A sign of recklessness, maybe the driver likes risks. Check him. A stalled engine? Nerves, maybe. Check him.[6]

The game is rather like a repeated game of matching pennies with one-sided strategic costs, played between customs inspectors on the one side, choosing which trucks to inspect, and a succession of hauliers on the other, deciding whether or not to put contraband in their trucks. The hauliers, though they act independently of one another, are reckoned by the inspectors to have similar characteristics, and the inspectors treat the succession of trucks as moves made by a single player, rather than by a succession of players. We could model this situation by a game in which there are just two players, one of whom, representing the hauliers, observes only a tiny fraction of the moves made by the player representing the inspectors. Thus, although the case of

[6]The passage, which is taken from [5], describes the efforts of customs inspectors at the port of Folkestone in England to detect smuggling activity by road hauliers.

a repeated game of matching pennies in which only one player faces strategic costs and the same player does not observe the moves made by the other player is a special case, it is an interesting special case.

A further comparison is worthy of note, between the analysis of games in which a computer plays against a human that can not observe the computer's actions, and the analysis of games in which a computer plays against a human that can observe the human's actions. If the human can not observe the computer's actions, then the problem facing the human is essentially to produce a sequence of h's and t's that looks random, or that they think will look random to the computer, subject to their strategic costs. If we collect a data set on human play against some computer player in a game in which the human does not observe the actions of the computer, then we can easily go back again and again to test alternative computer players against the same data. In other words, since the human did not observe the computer in the first place, and was presumably uninfluenced by it, we can use the data as a test set for new computer players that might be proposed. If, on the other hand, a game is played in which the human player does observe after each stage the actions played by the computer player, then we can not safely assume that the process by which human play proceeded was independent of the specification of the human player, and we can not usefully test the same data set against a different computer player - instead we would have to arrange a new experiment against a fresh human player.

3. Computation of strategic costs

Even if we settle the question of what strategy is appropriate to play in a repeated game of matching pennies against a player with any given strategic costs, there may remain several practical difficulties regarding the computation or estimation of a player's actual strategic costs. Although their exact nature and severity vary widely according to the model of strategic costs being used, we can, roughly speaking, describe at least two sorts of difficulty.

Suppose strategic costs of pure strategies are described as a function of the complexity, in some sense, of the strategies. For example, we might say, following Rubinstein [10], that the cost of using each pure strategy in a game is an increasing function of the smallest number of states that a Moore machine would have to possess in order to be able to carry out the strategy. The first sort of difficulty arises in evaluating the complexity of pure strategies. In some cases this difficulty is quite 'deep'. For example; if, following Smith [11], we define the appropriate sense of complexity of a strategy to be the Kolmogorov complexity of a coding of the sequence of action profiles implied by the strategy and the chosen pure strategies of other players, then we face the difficulty that Kolmogorov complexity is not generally computable, and can only generally be evaluated approximately. In other cases exact evaluation may be generally possible but nonetheless complex in practice. Suppose that the cost to player 1 of each pure strategy in a repeated game of matching pennies is a function of the complexity in some sense of the strategy, and that randomisation is costless. If player 1 may use a mixed strategy that results in the selection of each pure strategy with strictly positive probability, then in order to compute an appropriate strategy, player 2 may have to compute the complexity of each of 1's pure strategies. However, each player in a T-times repeated game of matching pennies may have as many as 2^{2^T} pure strategies, and even if the complexity of each one of these strategies could be calculated, it may not be practicable to do so. It may be possible to reduce the number of computations required, by arguing that above some known level of complexity, more complex strategies are prohibitively costly or even unavailable.[7] But this may not be enough, unless we are prepared to impose very strict limits on the complexity available to players. For example; even for quite modest N, the number of N state Moore machines that could be used to play strategies in a repeated game of matching pennies is large enough that it may pose practical computational difficulties. Exactly how many N state Moore machines are there that could be used to play strategies in a repeated game of matching pennies? Each state could be used to play one of two actions, and stemming from each state are two possible transitions, each of which might lead to any one of the N states. Finally, the start state might be any of the N states. Factoring, it is apparent that there are $N^{3N}2^N$ such machines. Of course, some of these machines effectively duplicate one another, since we have not ensured that each state could be reached on some play path. To be more precise, if k of the N states of a machine can never be reached by any transition from the start state, then the various formulations of those k states have no effect on the function of the machine, which is effectively just an $(N-k)$ state machine. Removing the multiple-counting, the number of non-equivalent N state Moore machines is

$$N^{3N}2^N - \sum_{k=1}^{N-1} (N-k)^{3(N-k)}\, 2^{(N-k)} \left[2^k N^{2k} - 1\right] . \tag{3.1}$$

This is still exponential in N, and very large for quite small values of N.[8] And it is not exceptional or particular to that sense of complexity for the number of pure strategies that are no more complex than some specified amount to be an exponential function of the amount of complexity. For a second example, a pure strategy with bounded full recall r is defined by Aumann & Sorin [2] to be one where the action played at any stage can depend only on the history of pure actions played in the previous r stages. There are $(2\times 2)^r$ possible histories of length r in a repeated game of matching pennies, and a pure strategy is a mapping from histories on to a pure action set. Each player has two available pure actions, so a player has

$$2^{(2\times 2)^r} = 2^{4^r} \tag{3.2}$$

pure strategies of recall r. Asymptotically, this expression increases even faster than 3.1.

[7] Similar reasoning is implicit in many evolutionary models of strategy selection dynamics. For example, it is common to model strategy selection in repeated plays of a game by means of a genetic algorithm, which imposes an upper limit on the complexity of strategies that might be selected.

[8] Some tabulated values of expression 3.1 are:

N	1	2	3	4
value of 3.1	2	242	152,466	263,226,650

A second sort of difficulty can be related back to the case, described already in section 2 of this paper, in which a player actually uses some pure strategy with probability close to but not exactly equal to one. If this might be the case then any Bayesian attempt to update a prior distribution on the pure strategy that has actually been selected by such a player must be conditioned on beliefs about the mixed strategies that he might use. Implicitly, it is not enough to know which pure strategy he is most likely to have selected, it is important also to know which other pure strategies might have been selected. Heuristically speaking, it is important to understand what sort of mistakes he is likely to make. Unfortunately, most models of strategic cost that can be found in the literature do not account for the making of mistakes with positive probability, so an acount may need to be provided. The importance of this point can be illustrated easily. For example, say we begin by assuming that strategic costs of each strategy are a function of the smallest number of internal states that a Moore machine must possess in order to carry out the strategy. Suppose that in a repeated game of matching pennies, in which a player, say player one, can only observe her own past actions, her sequence of actions in the first ten stages of the game is

$$hththththt\ ,$$

which could have been generated by a two state Moore machine. If her action in the eleventh stage is t, then, if we only consider non-probabilistic and deterministic Moore machines, her whole sequence of eleven actions could not be explained as the output of any Moore machine with less than eleven internal states.[9] However, what should the other player do, after the eleventh stage? He could assume that player one would not have made a mistake, and thus that player one must be using a machine with at least eleven states, and also accept any concomitant inference concerning the strategic costs of player one. Or, he could regard the internal states of player one's machine as playing each action other than their designed action with some small probability, in which case he may maintain a belief that player one is using a two state Moore machine. Or, he could assume that errors happen, and that they result in permanent mutation of player one's machine, in which case he might come to believe that player one's machine has mutated to one in which each of two internal states plays the action t. We could go on, but these three alternative accounts already amply demonstrate that the way in which errors in carrying out strategies are accounted for can have dramatic effects on the beliefs which players develop about one another's strategies.

4. An experimental study

We have identified a number of problems, some of them quite foundational, that stand in the way of inferring players' strategic costs from their play of penny-matching games. However, we have also argued that some of these problems can be 'worked around', at least in certain simple versions of the games. In this section we present the results of a short series of experiments with some of these versions.

[9]Note that if we allow the use of probabilistic Moore machines but still base their cost only on the number of their internal states, player one could simply choose a one state machine and thereby play the Nash strategy of the game without strategic costs, with the same payoff.

4.1. The experimental setup

Observations on human gameplay were gathered in two sequences of experiments, the sequences taking place on successive afternoons during the first week of June, 1999. Each experiment in the first sequence had a single human subject, who at the start of the experiment was presented with a piece of paper, with the following instructions printed on one side of it.

> Please read the following instructions carefully, then signal when you are ready to proceed. You can refer back to these instructions at any time during the experiment.
>
> Imagine a game between two players, who simultaneously call out either 'heads' or 'tails'. For one player, the object of the game is to make the same call as her or his opponent, and for the other player, the object is to make a different call. The actual calls must either be the same or different, so one player must win and the other must lose.
>
> Your task is to play this game one hundred times against a computer. You will win one point each time the calls match, and the computer will win if they differ. At each play of the game, the computer will be aware of the calls that both you and it made at each previous play of the game. You, however, will not be told any of the computer's calls until after the experiment is over.
>
> Write your calls down, in any pattern you like so long as the sequence is clear. Don't change any of your calls however, once you have written them down. For brevity, you can if you like just write 'h' for 'heads' and 't' for 'tails'. If you're not sure that you have written down a hundred calls, you can write a few more just in case. Writing extra calls won't affect your score, but if you write too few then the computer will win the remaining plays by default.
>
> Signal when you have finished. Your calls will then be fed into the computer: before each call is fed in, the computer will have to make its own call for that play of the game. You will learn your final score, which will be the sum of plays that you won less the sum that you lost, shortly afterwards.

The results were then gathered. Certain points regarding the experimental procedure are worthy of note. First, time pressures were minimised - instructions did not have to be commited to memory and there were no strict time bounds on completion. Second, no restriction was placed on the medium to which subjects wrote their calls. This, we thought, might give us evidence for two effects; firstly, that there might be some important difference between the responses of subjects who wrote into a computer and those who wrote on paper, and secondly, that the arrangement of a subject's written calls might hint at some pattern (that the subject may or may not be aware of) such as the subject writing down (and perhaps deciding upon) calls in groups of five or six, rather than individually. Third, subjects were told that they would not be penalised for writing too many responses. Our hope from this was to find that subjects generally wrote a few more calls than the required hundred, thereby providing some evidence that they were not very effectively ‘keeping count’ of the number of stages of the game that they had played. Fourth, subjects were not paid any money for their participation. An initial posted advertisement asking for volunteers produced a good number of responses, in addition to which our available funding would not have supported more than quite token payments. Appropriate payment formulae could easily be devised, however. In particular, subjects could be paid a flat fee for participation, plus some multiple of their eventual points score (which might of course be negative). Furthermore, it could then be quite revealing to vary the size of the multiple of their points score in this formula, since by varying the reward for winning we may encourage the subjects to use more or less sophisticated strategies. By this means, we could investigate the whole schedule of costs for different degrees of complexity of strategies, rather than just one point on the scale.

On the second day, another sequence of experiments was carried out, this time ostensibly between human players (each one of whom had also been a subject the previous day). Each of the pair of subjects was given the following instructions, printed on one side of a sheet of paper.

> Please read the following instructions carefully, then signal when you are ready to proceed. You can refer back to these instructions at any time during the experiment.
>
> Imagine a game between two players, who simultaneously call out either ’heads’ or ’tails’. For one player, the object of the game is to make the same call as her or his opponent, and for the other player, the object is to make a different call. The actual calls must either be the same or different, so one player must win and the other must lose.
>
> Your task is now to play this game one hundred times against a human opponent. The referee will tell you in advance if you are to be the player who always wins when the calls match, or the one who always wins when they differ.
>
> For each play, write your move in private then show it to the referee. The referee will write both calls and the name of the winner of that play on a piece of paper that all three of you will be able to see throughout the experiment. You may keep score privately if you wish.

Thus, each player thought that her or his task was to play the game against a human opponent, who was present with them, and whose past play was readily observable together with the score sheet.kept by the referee. We have, for interest, recorded the subjects’ calls against one another, but for reasons we have already discussed, we did not analyse their play against one another directly. Rather, we took each subject’s series of calls and played it against the computer players, as before. Our hope was to discover evidence that something, perhaps the prompts provided by awareness of the other human player’s calls, or stress induced by constant awareness of a running score, would either encourage subjects to generate a less predictable pattern of calls, or lead to them producing a more predictable pattern.

There remain a number of interesting tests that we could carry out with this data already collected. Obviously, we could create new computer players against which to play their calls (and to test against fresh subjects). In particular, we could search with more sophistication for evidence of interaction between the play of human players, by playing each player's series of moves independently against computer players, as we have done, but allowing the computer players to use the past calls of both human players as aids to predicting the calls of the player under study.

4.2. Description of the computer players

Each sequence of a hundred plays recorded by each experimental subject was played against each one of six computer players, christened $PROP$, $PREV$, $SEER2$, $SEER3$, $SEER4$, and $SEER5$. Algorithmic descriptions of the computer players are as follows.

$PROP$ operates as if on the assumption that the probability with which its opponent plays h at any stage is equal to one minus the relative frequency with which she or he has played h in past stages of the game (except that for the first stage of the game it operates as if that relative frequency of h is one half).[10] It plays its own best response to this belief. In other words, if the subject has played h in more than half of all past plays, then $PROP$ guesses that the subject is more likely than not to play t in the next stage, and so $PROP$ plays t. Why this should be a good strategy is hard to explain using any current theory of strategic costs. Insofar as it is a good strategy, it may indicate that human players are afraid of playing nonrandomly and use their relative frequency of h and t moves as an indicator of their randomness.

$PREV$ is quite straightforward. It operates as if on the assumption that its opponent will always play the same action as she or he did in the previous period, and plays a best response given that prediction. For its very first move, it plays h with probability one half. This can be justified quite easily in terms of the theory that strategic costs are a function of Moore machine complexity, as in [8]. Suppose strategic costs are such that a human player can not justify using any strategy that could not be implemented by some Moore machine with a single internal state, but that at each stage with some probability the human fails to remember the action he carried out at the previous stage. It follows straightforwardly that $PREV$ is a best reply in these circumstances.

$SEER2$ is a simplified version of Hagelbarger's SEER. If there is a clear short period cycle in its opponent's recent moves, $SEER2$ is designed to operate as if successfully predicting the cycle. Otherwise it plays as if randomly. For the first four stages, $SEER2$ always plays each action with probability one half. This player was

[10]Plays of h and t were coded for the computer players as 1's and 0's respectively. Whenever any of the computer players was required to play a mixed action, its move was simply coded as the probability with which the computer player was supposed to play h. The computer's expected payoff at each stage was then calculated by the formula

$$|(\text{coded value of subject's move}) - (\text{coded value of computer's move})|$$

and this expectation was recorded as the computer's payoff.

used by us in some experiments a year ago, and was included here for purposes of comparison.

$SEER3$ is quite similar to $PREV$. It operates as if on the assumption that its opponent will always play the same action as she or he did in the previous period, and plays a best response given that prediction. For its first two moves, it plays h with probability one half. This can be justified quite easily in terms of the theory that strategic costs are a function of Moore machine complexity, as in [8]. Suppose strategic costs are such that a human player can not justify using any strategy that could not be implemented by some Moore machine with at most two internal states, and that the human operates as if using some such Moore machine; but that at each stage, with some probability, each internal state of this Moore machine mutates and, afterwards, always plays a different action. Depending on the precise costs of Moore machines of different sizes, and on whether state mutations are correlated, $SEER3$ may be a best reply in these circumstances. Consider, if the subject does act as if using a Moore machine with two internal states, and makes use of both states, then because she can only condition on her own past play, one would expect her to play ate each stage what she played two stages before.

$SEER4$ is similar to $SEER3$ and $PREV$, except that it operates as if on the assumption that its opponent will always play the same action as she or he did three periods before, and plays a best response given that prediction (except that for its first three moves, it plays h with probability one half). Suppose strategic costs are such that a human player can not justify using any strategy that could not be implemented by some Moore machine with at most three internal states, and that the human operates as if using some such Moore machine; but that at each stage, with some probability, each internal state of this Moore machine mutates and, afterwards, always plays a different action. Depending on whether state mutations are correlated, $SEER4$ may be a best reply in these circumstances.

$SEER5$ is slightly different to $SEER3$. Like $SEER3$, it can be justified as a best response to an opponent who is playing as if with some two state Moore machine that occasionally mutates during use. Unlike $SEER3$, it predicts its opponent's next move on the basis that mutations of the Moore machine's state transition function are as likely as mutations in its internal states. It could predict relatively well during a period of play in which its opponent appears to switch back and forth between using a one state Moore machine and a two state machine.

The operation of each of $PREV$, $SEER2$, $SEER3$, $SEER4$ and $SEER5$ is illustrated by table 4.1. The first four columns list sixteen possible sequences that could comprise the last four actions played by the computer's opponent, at stages $t-4$ through $t-1$, and each subsequent column lists the action that the named computer player would predict that its opponent would make after observing each of these histories.

$t-4$	$t-3$	$t-2$	$t-1$	$PREV$	$SEER2$	$SEER3$	$SEER4$	$SEER5$
h	h	h	h	h	h	h	h	h
h	h	h	t	t	h	h	h	t
h	h	t	h	h	0.5	t	h	t
h	h	t	t	t	0.5	t	h	t
h	t	h	h	h	h	h	t	h
h	t	h	t	t	h	h	t	h
h	t	t	h	h	0.5	t	t	t
h	t	t	t	t	0.5	t	t	t
t	h	h	h	h	0.5	h	h	h
t	h	h	t	t	0.5	h	h	h
t	h	t	h	h	t	t	h	t
t	h	t	t	t	t	t	h	t
t	t	h	h	h	0.5	h	t	h
t	t	h	t	t	0.5	h	t	h
t	t	t	h	h	t	t	t	h
t	t	t	t	t	t	t	t	t

(4.1)

4.3. Interpretation

All except one subject eschewed the use of a computer and wrote their plays on blank paper. Every subject in the second sequence of experiments wrote three or four long columns or rows of plays, and each one of them failed to realise when they had made one hundred moves. Only 7 of the other subjects made the same mistake. 14 out of 24 subjects in the first sequence of experiments wrote their plays in ten columns of ten plays each, and none of these 14 wrote more than 100. A facsimile of one of these response sheets is reproduced in figure ??. The time taken by each subject to complete the experiment (including time taken to read the instructions) ranged between 9 minutes and 31 minutes, with an average of about 15 minutes for the single subject experiments and about 20 minutes for the twin subject experiments.

Some basic frequency information from the responses is given in table 4.2. Row A contains bin labels, which are inclusive. Row B gives the frequency with which subjects played h. Thus for example, there were 5 subjects who used h between 41 and 45 times (inclusive) out of 100. Row C gives the same information, but for just the responses from the single subject experiments. Row D gives the same information again, but for just the responses from the twin subject experiments. Row D gives the frequency with which subjects responses diplayed specified numbers of runs, that is of unbroken sequences of h's or t's. Rows E and F provide similar information, restricted to responses from the single and twin subject experiments respectively.

If subjects were in fact playing h at each play with independent probability of one half, then one would expect the distributions of the frequency data in rows B through C to approximate to normal distributions with mean of 50 and standard deviation of 5. As it happens the sample means were 48.7 for the whole sample, 48.9583 for the subsample of 24 single subject experiments, and 47.666 recurring for the subsample of 6 twin subject experiments. At a confidence level of, say, 97.5%, we can not reject the hypothesis any one of these figures is significantly less than 50, although they are certainly suggestive. As is easy to see from the table, the hypothesis that the frequency data is also distributed with standard deviation of 5 can not be rejected by a two-tailed test at a confidence level of 95%, either for the whole sample or either one of the subsamples.

The data on runs can be used to formulate a rough test of independence of subsequent plays, for if each play was indeed drawn with equal and independent probability then over 100 plays the distribution of runs should approximate to a normal distribution with known mean and variance. For each response, ie. each sequence of 100 plays, we used an ML estimator of the probability of any play being h, from which we deduced ML estimates of the mean and variance of the sampling distribution of the number of runs under the null hypothesis that plays were drawn with equal and independent probability. We then computed, for each sequence of 100 plays, the deviation from mean of the number of runs. Using one-tailed confidence intervals of 97.5%, we found that in the case of 6 of the 24 single subject responses and 1 of the six twin subject responses we could reject the null hypothesis, on the grounds that there were too many runs, whilst in the case of 1 of the 24 single subject responses and 2 of the six twin subject responses we could reject the null hypothesis on the grounds that there were too few runs. We also split each response into two subsequences, one comprising the first fifty plays and the other comprising the second fifty plays, and computed the sample correlation coefficient between the number of runs in the first subsequence and the number of runs in the following subsequence, in order to investigate whether individual subjects might tend systematically to play either too many or too few runs. The results of this were inconclusive however, the coefficient of correlation being just 0.5908. However, if we use a single ML estimate for all subjects of the probability of any play being h, we can reject, using a two-tailed 95% confidence interval, the null hypothesis that all plays were drawn with equal and independent probability. As can be seen in the table, the distribution of the number of runs is rather skewed, with many responses showing 'too many' runs.

The information represented in table 4.2 suggests that the subjects may have been playing randomly. But it also offers support to the hypothesis that subjects were rather trying to appear as though they were playing randomly, and using the relative frequency with which they were playing h, and the number of runs their play showed, as simple indicators of the apparent randomness of their play, especially since subjects on average showed slightly 'too many' runs - possibly to be sure that they were not interpreted as playing a 'simple' fixed strategy. This is useful, since it supports the hypothesis that subjects indeed understood the nature of the games they were asked to play. We can also derive some support for the hypothesis that they were attempting to play their Nash strategies, which if true could make the task of analysing experiments such as these much simpler.[11] We can investigate further

[11] One curious caveat: although the overall relative frequency of h was not significantly different to one half, the very first play by 24 out of the 30 subjects (including 21 out of the 24 subjects of single subject experiments) was h.

the hypotheses that subjects were using the relative frequency of h in their play and the number of runs that their play showed as indicators of the apparent randomness of their play. If the first one of these was true, then we might see the computer $PROP$ perform rather well, since it is designed to anticipate subjects bringing the relative frequency of h in their play back towards one half. If the second hypothesis was true, and if in addition subjects tended to put in slightly too many runs, then we would see the computer $PREV$ perform rather badly, since $PREV$ loses just when its opponent's play starts a new run. As we shall see, both predictions are borne out.

A	0-30	31-35	36-40	41-45	46-50	51-55	56-60	61-65	66-70	70-100
B	0	0	0	5	19	4	1	1	0	0
C	0	0	0	4	15	3	1	1	0	0
D	0	0	0	1	4	1	0	0	0	0
E	0	1	2	4	5	6	5	5	2	0
F	0	1	0	1	5	6	5	4	2	0
G	0	0	2	3	0	0	0	1	0	0

(4.2)

Tables 4.3, 4.4 and 4.5 show summary statistics of computer player performance against the responses made by; all subjects, just the subjects of single subject experiments, and just the subjects of twin subject experiments, respectively. The column headings F, G, H, I and J describe the same statistics in each case. In each table, column F shows the mean score achieved by each computer player against the subjects in the respective pool of subjects. If the computer player performs no better than a completely random player of h and t moves, then its expected score would be 50 and the variance of its score would be 25, and the distribution of its sample mean against a sample of n players would be approximately normal, with mean of 50 and variance of $\frac{25}{n}$. Two-tailed 95% confidence intervals (which can otherwise be interpreted as critical values for one-tailed tests at a 97.5% confidence level) to test this null hypothesis are; 50 ± 1.58833 for entries in table 4.3, 50 ± 1.98542 for table 4.4, and 50 ± 7.94167 for table 4.5. G shows the variance of the distribution of each computer player's scores against the subjects in the respective pool of subjects. H shows the coefficient of skewness of the distribution of each computer player's scores against the subjects in the respective pool. I shows the coefficient of kurtosis of the distribution of each computer player's scores against the subjects in the respective pool. Each instance of J is a statistic computed from the corresponding instances of H, I, and the number n of subjects in the respective pool, by the formula

$$J = n\left(\frac{H^2}{6} - \frac{(I-3)^2}{24}\right) .$$

This statistic is approximately distributed according to a χ^2 distribution with 2 degrees of freedom, and can be derived as an approximate Wald test of the null hypothesis of normality of the distribution from which the instances of H, I and n are taken. Using the 95% confidence limit of 5.99 for a χ^2 distribution with 2 degrees of freedom, we can accept the null hypothesis that the distribution of scores by each computer player is normal if the corresponding figure in column J is less than 5.99.

In the complete sample, represented in table 4.3, $PROP$ can be seen to have performed significantly better than a random player would have been expected to perform (winning over 53% of the available points), just as predicted. Also as predicted, $PREV$ performed significantly worse than a random player would have been expected to perform. Both conclusions also hold in the restricted sample made up of just the subjects of single subject experiments. However, neither conclusion holds in the restricted sample of subjects of the twin subject experiments. In that sample, $PROP$ performs very much as we would expect a random player to do, and $PREV$ does rather better, winning 55% of the available points (though this is not a large enough margin to be significant in this sample). This dichotomy can be related back to the distributions of runs in subjects' play. In the twin subject experiments, unlike in the other ones, subjects on average displayed rather fewer runs than a completely random player would have been expected to show, and hence guessing that one of these players would play the same action that she or he played last would be a rather good strategy. Quite why this should be the case is hard to discern. It may be that we were right to suppose that subjects use the relative frequency of h and the frequency of runs as indicators of randomness, but that subjects may use other indicators too, and when playing along with another subject several possible indicators are provided by the combination of their own sequence of plays and that of the other subject. We have not yet investigated any such connections statistically, but some suggestive evidence was provided by remarks made by subjects after the experiments ended. One subject, when asked how she had decided upon her plays, said that for about the first fifty, she tried to guess randomly and independently of what her fellow subject guessed, but that after that she tried to outguess her fellow subject. As it happened, during the first fifty plays her performance against the computer players was rather like that of one of the subjects of the single subject experiments, and very poor against her fellow subject. During the second fifty plays, her performance against the computer players was more typical of a subject of the twin subject experiments, and she was significantly more successful against her fellow subject.[12]

The performances of $SEER2$, $SEER3$, $SEER4$ and $SEER5$, all of which were quite similar in design to Hagelbarger's $SEER$, were rather disappointing. None of them performed significantly better than a random player would do on average, indeed only two of them performed even slightly better, and that only in the sample of subjects from twin subject experiments. $SEER2$, $SEER3$, and $SEER5$ did not per-

[12] After the experiments ended, several subjects made interesting claims. One pair of subjects knew each other already (the others did not) but although they were intensely competitive and unusually slow in their play, they were surprised when looking over their play after the experiment, saying that it looked very non-random. Perhaps this just indicates that they were unfamiliar with the appearance of random sequences (particularly regarding the frequency of runs) but there were several particular passages of play that they were surprised by. The most curious was a sequence in which one player played t ten times in succession, without beating his fellow subject a single time. The loser here gave a very complicated description of his behaviour, saying that each play was the outcome of a non-random decision. The winner said that he felt that the loser (who was trying to match guesses) was simply too stubborn to change while he was on a losing streak.

form significantly better or worse than an average random player. $SEER4$ performed significantly worse than that in the complete sample and the sample of subjects from single subject experiments, and worse (though not significantly worse) in the sample of subjects from twin subject experiments - which indicates something positive since it indicates that a machine which was programmed to do the opposite of what $SEER4$ would do would do rather well. We are not sure how best to explain the latter result. The poor performance of the other $SEER$-like machines was surprising, being significantly worse than both the performance reported by Hagelbarger for his own machine, in [6], and the results we obtained ourselves just a year ago. One possible reason for the disparity is that both Hagelbarger's machine and our own machine of last Summer demanded that subjects play their moves directly into the computer.[13] Almost all our subjects this year chose to write on paper, and it may be that it was less costly with this medium to refer to the external record of their decisions, and that this encouraged them to perform some checks for randomness that they would not have done had they been forced to input their moves directly to a computer. We can not check this hypothesis without a further experiment. Our one subject who did use a computer did in fact perform much worse than average against $SEER2$ and $SEER5$, the computers most like $SEER$ and our own computer from last year, but neither result would be statistically significant, and, moreover, this subject was not forced to write directly in to a computer.

Overall, we were quite surprised by many of the characteristics in practice of the performance of our computer players. We have attempted to derive some machines that would be best responses given beliefs about players' strategic costs that are consistent with certain models of strategic cost. Not only was our success with those machines limited, but they were difficult to derive, and correspond to only a very narrow part of the range of models of strategic cost that have currency in the literature. This suggests to us that the best direction for further experimental work is to maintain the informational asymmetry in play between human and computer players, since this makes it much cheaper to test alternative computer players, and to focus attention first on finding machines that perform well against human players and then explaining why they do well.

computer	F	G	H	I	J
$PROP$	53.2667	36.2456	0.2243	2.27575	0.90729
$SEER2$	49.2	33.5767	2.00937	9.25178	69.0438
$PREV$	47.8667	71.2989	0.26126	2.43162	0.74511
$SEER3$	50	65.0667	1.41792	7.24078	32.5327
$SEER4$	47.1333	38.4322	-0.592	2.61462	1.93783
$SEER5$	49.4333	37.1789	0.20741	2.49846	0.52953

(4.3)

[13] For our own experiments with such a computer, we ran one series of experiments in which subjects were informed of their running score total, and another series in which they were not. In both series, the computer performed significantly better than an average random player.

computer	F	G	H	I	J
$PROP$	53.875	35.651	0.15956	2.39202	0.47148
$SEER2$	49.0833	38.7847	2.0534	8.77609	50.229
$PREV$	46.0833	58.4931	0.6198	3.72645	2.06433
$SEER3$	49.5417	70.9149	1.75976	7.82112	35.6302
$SEER4$	47.6667	39.1389	-0.8368	3.0233	2.80178
$SEER5$	49	39.3333	0.30099	2.63577	0.49505

(4.4)

computer	F	G	H	I	J
$PROP$	50.8333	31.2222	0.5163	1.92076	0.55776
$SEER2$	49.6667	12.4722	0.17143	1.4706	0.61415
$PREV$	55	58.9167	-1.486	3.6906	2.32736
$SEER3$	51.8333	37.4722	-1.3271	3.39155	1.79963
$SEER4$	45	29.9167	-0.8353	2.53123	0.7526
$SEER5$	51.1667	24.8056	0.0709	1.26254	0.75972

(4.5)

5. Conclusions

We started with two objectives in mind. First, to develop an understanding of how human players actually approach games like matching pennies in which a traditional analysis suggests that they should use simple nondegenerate mixed strategies, and second, to develop an empirical critique of theories of strategic cost. We have argued that these objectives are linked - that human players actually play games like matching pennies quite 'rationally', given that they face certain strategic costs. It follows, if this is true, that we must develop an empirical critique of strategic cost theories in order to understand fully how humans actually approach games like matching pennies (as we have noted, there are other problems, concerning what it means for players to be 'rational'). Also, we could look at what actually happens in games like matching pennies and reason back to the patterns of strategic costs that could have given rise to such play, thereby achieving an empirical critique of strategic cost theories. Hopefully, this paper represents a small advance towards each of these objectives.

References

[1] D. Abreu and A. Rubinstein, The structure of Nash Equilibrium in Repeated Games with Finite Automata, *Econometrica* **56** (1988), 1259-1281.

[2] R. Aumann and S. Sorin, Cooperation and Bounded Recall, *Games and Economic Behavior* **1** (1989), 5-39.

[3] R. Axelrod, "The Evolution of Cooperation," Basic Books, New York, 1984.

[4] M. Bacharach, H. Shin and M. Williams, Sophisticated Bounded Agents Play the Repeated Dilemma, Applied Economics Discussion Paper no. 143 (1992), Institute of Economics and Statistics, University of Oxford.

[5] R. Carroll, Highway robbery, *The Guardian Newspaper*, London, Section 2 (18 May 1999), 2.

[6] D. Hagelbarger, SEER, A SEquence Extraction Robot, *I.R.E. Trans. on Electronic Computers* **EC-5** (March 1956).

[7] N. Megiddo and A. Wigderson, On play by means of computing machines, IBM Research Report RJ 4984 (52161), (1986), IBM Almaden Research Center, San Jose.

[8] A. Neyman, Bounded Complexity Justifies Cooperation in the Finitely Repeated Prisoner's Dilemma, *Econ. Letters* **19** (1985), 227-230.

[9] R. Radner, Collusive Behavior in Noncooperative Epsilon-equilibria of Oligopolies with Long but Finite Lives, *J. Econ. Theory* **22** (1980), 136-154.

[10] A. Rubinstein, Finite Automata play the Repeated Prisoner's Dilemma, *J. Econ. Theory* **39** (1986), 83-96

[11] J. Smith, On Kolmogorov complexity and the costs of carrying out strategies, Applied Economics Discussion Paper 214 (1999), Institute of Economics and Statistics, University of Oxford.

[12] J. Von Neumann and O. Morgenstern, "Theory of Games and Economic Behavior," third edition, Princeton University Press, Princeton, 1953.