

# Demonstrating the Possibility of Pareto Inferior Nash Equilibria

Nicolaas J. Vriend

One of the serious challenges when teaching game theory to undergraduate students is to explain the occurrence of a multiplicity of equilibria and the related problem of equilibrium selection. Many equilibria are simply considered to be totally “unreasonable” and “senseless” by many students and worth their attention only because they fit the technical definition of an equilibrium as given in the textbook, which implies it might be useful to know how to find these equilibria only because of the final exam.

A particular example of such “absurd” equilibria is the occurrence of Pareto-dominated equilibria, as in the bimatrix game Matching Colors (Figure 1). The (red, red) equilibrium with a payoff of 100 to each payer is not considered to be a serious possibility by most students. Even when they understand how it fits the definition of a Nash equilibrium (NE), they believe only exceptionally boundedly rational people would fall for it. (In some sense, it seems that many students are *kibitzers* [Binmore 1992, 397], people who watch games but do not play themselves and who tend to believe they know the right moves better than the players themselves.) Possible explanations referring to historical accidents and the dynamics of social norms appear somewhat abstract and receive mainly skeptical reactions in the classroom. This seems a typical situation in which a class-

**FIGURE 1**  
**A Bimatrix Game: Matching Colors**

		Player 2	
		Red	Blue
Player 1	Red	100	80
	Blue	100	250
		Red	Blue
		100	250
		80	250

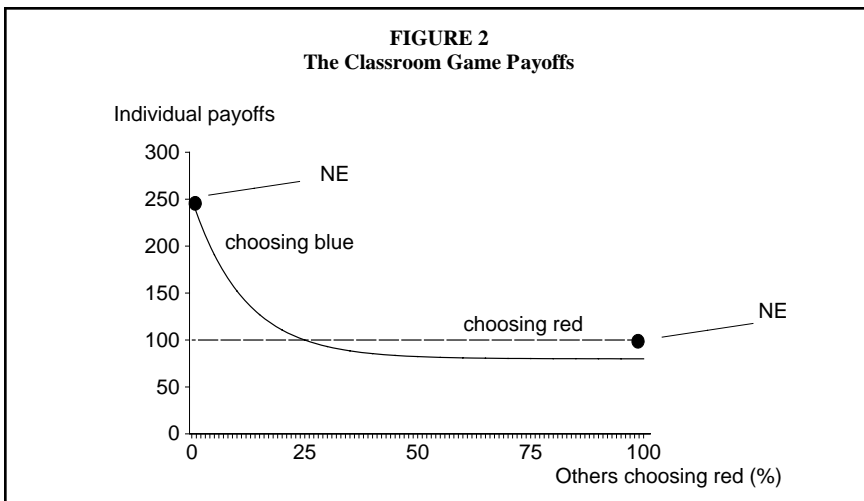
*Nicolaas J. Vriend* is a Reader of Microeconomics at Queen Mary and Westfield College, University of London, Department of Economics (e-mail: n.vriend@qmw.ac.uk). The author wishes to thank two anonymous referees for helpful comments.

room game could teach them a “hard” lesson. If one can actually demonstrate to the students that such equilibria may occur quite naturally, they will without doubt start taking them more seriously.

It should be stressed that this article does not fit into the abundant literature on equilibrium refinements. New equilibrium selection arguments will not be presented nor will old arguments be tested. The sole objective is an educational one: presenting a teaching tool that might be useful to show students the plausibility of Pareto inferior equilibria.

### THE GAME ITSELF

We use the  $n$ -person game shown in Figure 2, of which the presentation is based on related games in Schelling (1978). All individual players in a group choose simultaneously between two strategies (red or blue). The individual payoff corresponding to red is independent of the number of other players choosing these strategies, but the payoff to a player choosing blue is a decreasing function of the number of other players choosing red. The critical number in this game is 25 percent. If more than 25 percent of the other players choose red, then choosing red is strictly better, whereas if less than 25 percent choose red, players choosing blue are strictly better off. Notice that there are two Nash equilibria (in pure strategies), with one Pareto dominating the other; all red (with a payoff of 100) and all blue (with a payoff of 250). The intersection itself is not an NE. Suppose 25 percent of the other people would choose red. Then you are indifferent. If you choose red, this would leave other people opting for red indifferent between red and blue, but other people going for blue will prefer to deviate toward red. Similarly, if you choose blue, other players going for blue will remain indifferent, but those who planned to choose red will deviate and choose blue.



## ORGANIZING THE GAME IN THE CLASSROOM

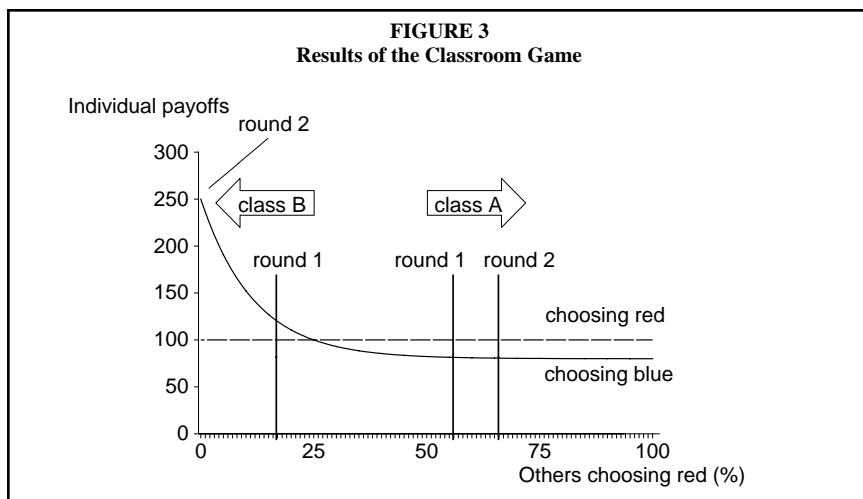
The game was organized in the framework of a game theory course for third-year undergraduate students in economics at the University of London, Queen Mary and Westfield College in October 1998. The course was organized with a weekly two-hour lecture for about 60 students, combined with two parallel one-hour classes of about 30 students each. We played two rounds of this game during the two parallel classes. Before the first round (in the second week of the term), the game was introduced providing a background story about chattering students (choosing red) who get a constant payoff of 100 independent of what other people do, whereas the students who pay attention to their teacher (choosing blue) suffer from an externality: When nobody chatters, it gives a high payoff of 250, but as more people chatter, this payoff falls quickly below that of students chattering. The exact payoffs available were explained by showing a graph on the overhead projector (Figure 2) but without indicating the two Nash equilibria. Also, no further analysis or explanation of the equilibria of the game was given. The number of rounds was not announced in advance. Moreover, no special attention was given to this game in class, and hence no elaborate coordination attempt was to be expected during the week between rounds 1 and 2. Notice that although this is a repeated game, the influence of an individual player on the final outcome is negligible anyway, as in a competitive market. The second round was played one week later in the two parallel classes. For each class, the students in round 2 were in principle the same as in round 1. The fact that the numbers in rounds 1 and 2 were not exactly the same in both classes (Table 1) reflects natural fluctuations in class attendance. This implies that a handful of students played only one of the two rounds. But what is important is that in each class the students were given the results of the previous round of that class, and of that class *alone*, indicating the outcome in the graph. Because the two parallel classes were back-to-back, the students of one class had no time to talk with those of the other class about their results.

Incentives for the students to seek to do the best they could were provided in the form of credit points. The weighted average of the payoffs realized in this game plus a series of other games during the term counted for 5 percent of the final credit of this course. The distinct dynamics in the two groups discussed below suggest that these incentives did work, inducing the students to a definite

**TABLE 1**  
**Results of the Classroom Game**

Class	First round	Second round
Class A		
Red	21 (55%)	22 (65%)
Blue	17	12
Class B		
Red	4 (16%)	0 (0%)
Blue	21	28

**FIGURE 3**  
**Results of the Classroom Game**



and nonerratic course of action. Nevertheless, readers who might feel somewhat uncomfortable about any scientific significance attached to results generated by such weak incentives are perfectly right. But they should bear in mind that there is an essential difference between laboratory experiments for scientific purposes on the one hand, and classroom games for educational purposes.

## RESULTS

Table 1 and Figure 3 present the choices made by the students for each class and each round separately. By chance, or any other unclear cause, class A happened to pick a point to the right of the critical 25 percent level, whereas class B selected a point to the left in the first round. This difference was amplified in the second round, with class B jumping immediately to the Nash equilibrium (all blue), whereas class A moved further into the direction of the other Nash equilibrium (all red).

## DISCUSSION

After collecting all results, they were presented and discussed in the central lecture with all students together. First, we analyzed the game with its Nash equilibria and then showed the results for class A. Although no complete convergence to any Nash equilibrium took place in that class, the students agreed that a couple of extra rounds would lead everybody in that class to choose red, the Pareto-dominated equilibrium. Showing then the entirely different results for class B impressed most of them. Having seen the dynamics of the outcomes, all students were convinced that they would stay at the respective equilibria forever. Obviously, whether these views would turn out to be actually correct in the long run is an entirely different question. It might be that there exist dynamic processes leading to different results in the long run (see e.g., Mailath [1998] for a survey

on evolutionary dynamics and Nash equilibria). But the essential point here is an educational one: convincing the students of the plausibility of inferior equilibria. And, looking at the views expressed by the students after the results were presented, that goal apparently was reached.

Notice that a key characteristic of this game is that it is a group game. This implies that *everybody* became an accomplice, and the students could not simply point to some other players' bounded rationality as an explanation for the outcome. This factor makes it very different from showing the occasional occurrence of Pareto-dominated equilibria in a series of one-shot two-player coordination games, where students could more easily argue that such outcomes are singular events, basically caused by random factors like the "mistakes" of a small number of individuals.

The results of the classroom game presented here seem most interesting whenever at least one class converges to the Pareto-dominated equilibrium. With a basin of attraction of 75 percent, there is a sporting chance this will happen. Perhaps even more instructive is when two parallel classes converge to a different equilibrium. Because the smaller basin of attraction is compensated by the higher payoff, this will also occur with a considerate probability. Related to this, if the course is taught to only one group, then it might make sense to split the group into two subgroups, although this obviously requires some additional administration. But what to do if no class converges to the Pareto-dominated equilibrium? The underlying idea of the game is that the more the students themselves are involved in generating the inferior equilibria, the more convincingly their plausibility is shown. Therefore, the fall-back options are, in order of preference, the results of a previous year of the same course, the results of the course of a colleague at the same university, and finally, the results reported in this article.

As the teacher of these classes, I was not surprised by these results. Right from the beginning of the term, the general climate in two parallel classes had evolved in a quite different way, in some sense corresponding to the nature of the two equilibria in the game. Many teachers must have experienced this phenomenon of a dissimilar evolution of social norms in different parallel classes during a term. But then, the point of this classroom game is to teach this to the students.

#### REFERENCES

- Binmore, K. G. 1992. *Fun and games. A text on game theory*. Lexington, Mass.:Heath.
- Mailath, G. J. 1998. Do people play Nash equilibrium? Lessons from evolutionary game theory. *Journal of Economic Literature* 36 (September): 1347-74.
- Schelling, T. C. 1978. *Micromotives and macrobehavior*. New York: Norton.