International Journal of Forecasting 18 (2002) 375–382
www.elsevier.com/locate/ijforecast

Behavioral decision making, forecasting, game theory, and role-play

Hersh Shefrin

Department of Finance, Santa Clara University, Santa Clara, CA 95053, USA

Abstract

Green's finding that the outcome of role-play provides forecasts that are superior to those of game theorists highlights some of the unrealistic assumptions used in traditional game theory. In this commentary I discuss how elements studied in the behavioral decision literature impact the manner in which people behave in conflict situations studied by Green, and in the spectrum auction conducted in the United States. The main behavioral elements discussed are loss aversion, myopia, and the winner's curse.

Green (2002) presents compelling evidence that when it comes to forecasting the outcomes of real world conflicts, role-play can be more effective than the forecasts made by game theorists. Yet most of the commentaries on Green's paper urge caution when interpreting his findings to imply that game theory should be abandoned as a forecasting tool. For example, Bolton (2002) points out that it is now common practice in business school negotiation classes to use role-play and game theory together. He suggests that in the future the most useful lessons will come from work that combines game theory with experimental economics, along the lines described in Erev, Roth, Slonim, and Barron (2002).

Although I concur with Bolton's general conclusion, I would argue that game theory's value as a forecasting tool will continue to be limited until game theorists revamp its structure along behavioral lines. Traditional game theory assumes that players are fully rational in respect to preferences (expected utility), judgments, and strategic choices. Yet the behavioral decision literature documents that preferences routinely violate the axioms of expected utility, and that judgments exhibit systematic biases and errors.

In making the case for the explicit incorporation of behavioral concepts into game theory, I draw on two sets of experiences. The first is my own experience from a setting described by Bolton, teaching a graduate business school course that makes dual use of game theory and role-play. The second is the experience of the spectrum auction being conducted by the US government.

1. Lessons from the classroom

One of Green's conflict situations is the '55% Pay Plan', a revenue sharing dispute between

E-mail address: hshefrin@mailer.scu.edu (H. Shefrin).

0169-2070/02/$ – see front matter
PII: S0169-2070(02)00021-3

US football players and team owners in the National Football League. Game theorists were asked to forecast if there would be a strike, and if so whether its duration would be short, medium, or long. The actual outcome of the dispute was a long strike. Among the game theorists' predictions of the actual outcome, 29% were accurate, a little better than the 25% associated with chance. In contrast, the corresponding accuracy rate from role-play was 60%.

What accounts for the apparent forecasting superiority of role-play in the '55% Pay Plan'? Is there anything missing in traditional game theory that would have led game theorists to underpredict the likelihood of a long strike? I suggest that several behavioral elements are missing from the game theoretic framework, elements that can be understood through the well-known game 'Dollar Auction'. In Dollar Auction, players bid for a $1 bill in an open cry ascending (English auction), and the winner receives the $1 and pays his or her last bid. However, the game has a twist. The second highest bidder also pays his or her last bid, but receives nothing.

Raiffa (1982) describes having induced two of his colleagues at Harvard Business School to play the game. Notably, the bidding began below $1. When it reached $1, the second highest bidder at the time realized that he could enter a bid for $1.01; paying $1.01 for $1 is better than losing the auction and paying 99 cents. However, the bidding did not stop at $1.01, and both players continued to bid. Indeed, by the time the bidding reached $3.10, emotional temperatures had risen, and the game showed no sign of concluding. Raiffa then intervened and terminated play.

Raiffa (1982) emphasizes that Dollar Auction is highly relevant to labor disputes such as the '55% Pay Plan'. A key lesson from Dollar Auction is that once both parties have incurred costs, there are psychological forces at work that tend to induce escalation in disputes such as the '55% Pay Plan'. In my own class, students play Dollar Auction for a $20 bill. In 7 years of play, the bidding has never stopped below $20, and only once did it stop at $20. Final bids ranged from $20 to $66, with the most frequent final bid around $30.

In his commentary, Goodwin (2001) points out that Green did not test role-play against game theory, per se, but against the forecasts of game theorists. However, with games such as Dollar Auction, it is possible to test predictions stemming from game theory. In the traditional game theoretic approach, players are assumed to prefer more money to less, and to exhibit Nash equilibrium behavior.

When players bid more than $1 to receive $1 in Dollar Auction, the outcome of the game is not a Nash equilibrium. Because people routinely bid more than $1 in Dollar Auction, it is natural to ask what factors drive their behavior? I suggest that the answer involves a mixture of behavioral phenomena that are familiar to readers of this journal:

- bounded rationality
- loss aversion leading to the escalation of commitment, and
- players' overvaluation of their own positions.

Bounded rationality leads most players to be myopic, and prevents them from thinking far enough out along the game tree to determine the Nash equilibria of the game. Loss aversion leads players to be willing to accept actuarially unfair bets in an effort to avoid loss. Kahneman and Tversky (1979) document that people tend to exhibit risk-seeking behavior in the domain of losses. This includes the loss of ego or pride as well as financial loss. In Dollar Auction, ego enters as an issue once the bidding exceeds the $1 being auctioned, and players attach increasing importance to 'winning' the $1 bill. The ego effect can be compounded by overvaluation, when a player misestimates the costs he or she incurs from escalating, relative to the other party.

These behavioral phenomena are quite pervasive, and affect the degree of bidding escalation in Dollar Auction, the length of the strike in Green's '55% Pay Plan' conflict, and the behavior of Argentina and the United Kingdom in the Falklands/Malvinas War, a conflict mentioned in Armstrong (2001). In the latter conflict, the Argentine military appears to have miscalculated the degree to which the United Kingdom was willing to incur the cost of a military campaign in order to recover possession of an isolated territory located far from its shores that holds little economic value or strategic value. Did the Argentine generals underestimate the impact of their actions on British pride?

Prospect theory, the psychologically based theory of choice developed by Kahneman and Tversky (1979), emphasizes the importance of 'framing'. In supplemental material available online, Green reports that the outcome to the conflict situation 'Nurses' Dispute' featured framing, in that 'the agreement was framed as a 7% pay increase' even though 'the dollar value of the agreement was half way between the stated positions of the two parties'. Interestingly, 'Nurses' Dispute' was one of two situations where unaided judgment outperformed the predictions of game theorists.

A prominent feature of most behavioral studies is that people display considerable variability in their choices and judgments. Notably, Raiffa (1982) reports that some of the role-plays he conducted featured considerable variability across groups. The variability issue also arises in Green's study. Green notes that the forecasting accuracy rate of role-play varies from 29% (Artists' Protest) to 82% (Nurses' Dispute), with the mean accuracy rate being 64%. Such variability heightens the need to focus on forecast distributions, as opposed to just expected outcomes.

2. Spectrum auctions

In 1994 the US government began to auction licenses to use portions of the electromagnetic spectrum. This auction was historic. Not only was it the largest in history, it provided the most intensive application of game theory ever. John McMillan (1994) describes how the players in this auction—from the Federal Communications Commission (FCC) who ran the auction to the large telecommunications firms who participated as bidders—all hired game theorists as consultants. What better test for a theory whose equilibria feature rational predictions and rational behavior, than a high stakes game where game theorists advise the major players. McMillan (1994) writes: "When the theorists met the policy-makers, concepts like Bayes–Nash equilibrium, incentive-compatibility constraints, and order-statistic theorems came to be discussed in the corridors of power" (p. 146).

McMillan wrote two important articles about the spectrum auction, one before the auction began, and a second (co-authored with Preston McAfee) in the middle of the auction. Because of the timing, the two articles together provide wonderful insights about applied game theory and prediction. See also the comments about the spectrum auction in Erev et al. (2002).

McMillan (1994) reports that there were lessons to be absorbed from other spectrum auctions, lessons that nicely illustrate how wide the gap can be between theory and practice. New Zealand had conducted an auction in 1990, using a second-price sealed bid auction on the advice of a consulting firm. Revenues turned out to be much lower than expected, perhaps because concerns about 'the winner's curse' led to very conservative bidding behavior.

Winner's curse is a phenomenon in which the winner of an auction overpays, because of a behavioral error. I introduce the concept of winner's curse to my students by auctioning off a jar of coins. The jar is transparent, but the sheer number of coins makes it difficult to estimate accurately the amount of money in the jar. Students are given time to inspect the jar, but are not permitted to open it. It is possible that the average estimate in the class provides an unbiased estimate of how much money is actually in the jar. However, the auction selects the high bid, not the average bid. That is why the winner of an auction is vulnerable to experiencing winner's curse. Typically the winner of my jar of coins auction pays between 5 and 20% more for the jar of coins than the value of the coins contained therein. Those who bid rationally should adjust their bids downward by an amount that factors in the exposure to winner's curse.

Like New Zealand, Australia too learned some lessons the hard way. Their auction, for two licenses, featured no penalty for default. This flaw left their auction ripe for being gamed by two dark horse bidders, Ucom Pty. Ltd. and Hi Vision, who outmaneuvered a consortium that included Rupert Murdoch and Telecom Australia. The two dark horses placed a series of bids, sequenced from low to very high, and then proceeded to default on their winning bids, only to replace themselves as the winning bidder with a lower winning bid. Ultimately Ucom succeeded in winning both licenses, which it then sold to other firms.

The FCC structured the US licenses so that each license assigned rights to a combination of frequency band and geographic location. Some locations were regional and others were national. Because the structure of the licenses was very complex, the FCC planned to hold five auctions, spread out in time, instead of one single auction.

The FCC faced a daunting task—how to structure an auction that allowed for efficient aggregation of licenses by bidders, discouraged bidders from being overly conservative in their bids for fear of succumbing to the winner's curse, and avoided past pitfalls such as the absence of default penalties and reserve bids. Moreover, although raising revenue was not the primary stated objective of the auction, the Clinton administration budget planners had publicly stated that they expected the auctions to yield up to $10 billion. This put considerable political pressure on the FCC.

The first auction took place in July 1994, for nationwide narrowband licenses. A second auction for regional narrowband licenses followed in October, just as the Nobel prize in economics was being awarded to the developers of game theory, John Nash, John Harsanyi, and Reinhard Selten.

Then in December, the FCC began its first major auction, for broadband licenses in major trading areas (MTAs). In this respect, the FCC had divided the US into 51 major trading areas (MTAs), with two 30-MHz blocks of spectrum associated to each. The FCC also established licenses for subdivisions of the MTAs called basic trading areas (BTAs), and associated portions of the spectrum. A fourth and fifth auction for BTAs was scheduled to follow the MTA-auction.

What made the US spectrum auction especially complex was that bidders were interested in developing regional and national networks. These bidders, therefore, were interested in aggregating licenses, with the value to them of a particular license dependent on what other licenses they were able to obtain. McMillan describes that concerns for efficient aggregation led the FCC to adopt a design based upon a simultaneous multiple-round auction.

McMillan (1994) concludes: "The FCC's adoption of a simultaneous multiple-round auction ahead of a sequential bid or a single-round auction—which are more conventional but arguably less effective for selling spectrum licenses—was a triumph for game theory. The intriguing next step will be to appraise its performance" (p. 160). So how did things turn out, both on the aggregation front and on the revenue front?

McAfee and McMillan (1996) suggest that in respect to the MTA auction, ‘‘the auctions facilitated license aggregation’’ (p. 167). They point out that PacTel, a major bidder, won the aggregation that it made no secret it was seeking, and that the other bidders came close to filling the gaps in their cellular holdings. As to whether the auctions had put the licenses into the hands of the right firms, McAfee and McMillan indicate that in 1996 it was too early to tell. However, they do note that if a large number of licenses were to be resold, then that would be evidence of an inefficient allocation.

How well did the auction work as far as revenue was concerned? Did the auction design serve to allay bidder's concerns about winner's curse? Recall that the FCC was under pressure to raise revenue, and hence worked to produce an auction design that would mitigate conservative bidding strategies. The initial revenue projection for the entire spectrum auction was $10 billion. Interestingly, the chairman of MCI, Bert Roberts, is quoted as having said: ‘‘The government is smoking something to think they are going to get $10 billion for these licenses’’.

Were bidders overly conservative? The two narrowband auctions raised over $1 billion. The MTA auction raised an astonishing $7.7 billion. McAfee and McMillan (1996) report that the average price per person covered by a license (price per pop) in the MTA auction was $15.51. The highest price per pop was in Chicago ($31.90), with Atlanta (at $28.58) and Seattle (at $27.79) also high. Interestingly, the Chicago price per pop was almost double that of New York and 20% higher than that of Los Angeles.

Writing just after the MTA auction, Andrew Seybold (1995) suggested that the high prices were indicative of winner's curse, if not mass undue optimism. He pointed out that Sprint paid about $26 per pop for the Los Angeles–San Diego area. Seybold then went on to adjust this price to reflect other factors. Beginning with the cost of moving existing microwave users off this spectrum, he added a build-out cost for infrastructure, assumed that only one-in-five would subscribe to the new services, and then factored in competition. All of this led him to suggest that the price paid per actual customer was closer to $1715. With a projected revenue stream of $600 per year per customer and a 20% profit margin, Seybold estimated that the payback period for the spectrum investment would be over 14 years.

That was in 1995. In an article for *Forbes* Magazine (September 17, 2001), Woolley assesses the profitability of communications firms, stating:

> The effects of so much competition have also been ruinous for US carriers. The six national providers—AT&T, Sprint, Verizon, Cingular, Nextel and VoiceStream—together ran up $3.5 billion in losses last year . . . For years the industry's unending flow of red ink was largely ignored by investors as a normal symptom for an industry in its adolescence. But the cellular business is two decades old now, and the industry's plummeting stocks and stalled public offerings seem to indicate that investors' appetite for funding losses is waning.

How difficult is it to forecast the outcome of these types of auctions with reasonable accuracy? To address this question, consider how the remainder of the spectrum auction has unfolded. The auctions for narrowband and MTA auctions described above were called the A and B block auctions. The next auction, for BTAs, was called the C block auction. The C block auction produced a fascinating outcome, which shares similar traits to the Ucom-outcome in the Australian spectrum auction.

Legal issues caused delays in holding the C block PCS (personal communication services) auction, but the auction began on 18 December 1995. This auction concluded on 6 May 1996. Altogether 493 licenses were sold for $9.2 billion, thereby taking total revenues well above the $10 billion estimate.

A San Diego-based firm, NextWave, was the auction's high bidder, placing $4.2 billion worth of bids for licenses covering 56 basic trading areas (BTAs). Notably, 10 days after the auction closed, licensees were required to pay 5% of their bids. NextWave made a payment of $130.8 million to the FCC. That, combined with the company's pre-auction deposit of $79.2 million, fulfilled the FCC's first down payment requirement. However, there were key defaults by other firms, and as a result, on 3 July 1996 the FCC began a re-auction of 18 defaulted C block licenses. In the re-auction, NextWave placed 17 out of 18 high bids in the first round. By 16 July 1996 when the C block auction closed, NextWave had purchased $4.8 billion of licenses that represented the third largest block of wireless spectrum in the United States including rights to all top 10 markets, 28 of the top 30 markets, and 40 of the top 50 markets.

After making the initial set of payments, NextWave owed the FCC over $4 billion, a staggering amount for a company with just 600 employees. In June 1998, NextWave found itself unable to fulfill its contractual obligations to the FCC. In consequence it filed for bankruptcy protection under Chapter 11. NextWave then spent more than $40 million on legal bills while trying to stay afloat and keep the licenses. It downsized to a skeleton crew of lawyers, engineers and marketing managers. It moved its headquarters from San Diego to a bare-bones hub in an industrial park in Hawthorne, NY, in order to be closer to Wall Street.

Eighteen months later, on 12 January 2000, the FCC announced that it was canceling NextWave's licenses and would re-auction them in July 2000. But on the very next day, NextWave announced that it would challenge the FCC's move. In February 2000, a bankruptcy court rejected the FCC's cancellation of NextWave's licenses. However, days later a Federal appeals court overturned the bankruptcy court's decision. So, in January 2001, the FCC completed a re-auction of the licenses for which NextWave had originally paid $4.8 billion. In the re-auction, these licenses sold for $16 billion.

On 22 June 2001, a three-judge panel of the US Court of Appeals for the District of Columbia Circuit said the FCC violated federal bankruptcy law by repossessing the NextWave licenses. This means that the court regarded NextWave to be the legal owner of licenses that other firms judged to be worth $16 billion.

In January 2001, the US government thought it had secured an additional $16 billion for the Treasury. It is reluctant to give this revenue up. NextWave argues that it stands ready to take the licenses, pay the $4.7 billion it originally bid for them, and put them to use. In fact, NextWave announced its decision to hire Lucent Technologies to build a 2.5-G network across its 95 markets.

The firms who won these licenses in the re-auction are eager to begin developing new products and services. They contend that NextWave lacks the technical expertise to carry out the projects, and have proposed a plan that would preserve approximately $12 billion for the US Treasury. These firms are offering to pay a certain portion of their bid to NextWave in order to settle the issue and take back the licenses. The firms suggested that NextWave receive $4 billion to $5 billion, with the government reserving the right to argue a lower number.

Game theorists have developed sophisticated techniques for analyzing multi-party conflicts such as the one just described. Could such techniques provide a reasonable forecast of the outcome? My sense is that game theory can provide a useful starting point. However, if history offers us any guide, it is that behavioral elements will prevent the parties from achieving an efficient outcome.

As this article goes to press in May 2002, the interested parties have not reached agreement about how to allocate the licenses, despite the large opportunity costs. The legal case is slated to go before the U.S. Supreme Court, but a decision is not expected for a year. In February, NextWave announced that it has activated its wireless network in 60 markets. Some carriers continue to negotiate with NextWave, while others claim to be withdrawing. In May, Verizon Wireless, which in value terms won roughly half the licenses in the January 2001 auction, announced that it has no interest in reaching a settlement with NextWave. Verizon indicated that it has reassessed what the licenses are worth, given the decline in market conditions. Notably, the FCC announced that it would refund 85% of the down payments from the controversial re-auction, although lawmakers have proposed that the FCC refund 100% of the down payments. Lawmakers also voted to delay the next auction, scheduled to take place on June 19. Opportunity costs continue to mount.

The U.S. spectrum auction has thus far raised about $20 billion, at least in theory. A similar auction in the United Kingdom took place in April 2000 and raised $34 billion. Although the two auctions were different in several respects, they shared a key feature—undue optimism leading to overpayment, the winner's curse. Writing about the winners and losers in the British auction, for the *Times of London*, Kaletsky (2002) states: ''The only losers will be stock market investors who were foolish enough to buy phone shares at the peak of the technology bubble—plus the financial analysts who goaded on corporate executives to bid vast sums for the rights to an untested technology of questionable value.''

Notably, spectrum auctions that subsequently took place in Europe—the Netherlands, Italy, and Switzerland—generated revenues that were disappointlngly low. Binmore and Klemperer (2002) attribute much of the disappointment to inappropriate auction design, rather than pared down expectations after the mania of the late 1990s. My own view is that the role of expectations has been critical, but one thing is clear: be it inappropriate expectations or inappropriate auction design, psychological factors have been key.

3. Conclusion

What is the practical relevance of game theory for the kind of day-to-day forecasting situations faced by practitioners, such as the individual conflicts studied by Armstrong (2002) and Green (2002)? It seems to me that Green is essentially correct, that in a head-to-head contest with role-play, traditional game theory comes out a distinct second.

What gives role-play its edge over game theory? In my opinion, the answer is that role-play outcomes emerge from the actual interaction of real human beings, whereas game theoretic outcomes emerge from the theoretical interaction of idealized human beings. The behavior of real human beings reflects emotion and errors in judgment, as embodied within psychological phenomena such as loss aversion, myopia, and winner's curse. In contrast, the idealized human beings modeled by game theory are rational in respect to their preferences, judgments, and choice of strategies.

The development of behavioral game theory is well underway (see Thaler, 1988; Rabin, 1994; and especially Camerer, 2001). The introduction of learning, as discussed by Erev et al. (2002) is a small, but important step. The predictive power of game theory will improve as the underlying assumptions and equilibrium concepts are made more realistic. Indeed role-play exercises and experiments should help game theorists to develop well-structured behaviorally based models.

McAfee and McMillan (1996) state that ''the real value of theory is in developing intuition'' rather than applying ''complicated models that try to capture a lot of reality'' (p. 172). It seems safe to say that from the perspective of the FCC, game theory was a powerful tool to have used in designing the US spectrum auction. The auction form they used provided a decent mechanism for bidders to aggregate licenses, and it raised more revenue than was forecast.

At the same time, McAfee and McMillan make a stronger assertion, stating that the role of theory ''is to show how people behave in various circumstances'' (p. 172). Of course, this is exactly the forecasting issue studied by Green. On this dimension, the evidence suggests game theory has had limited success. Winning bidders in the spectrum auction were afflicted by the winner's curse. And the C block auction produced a rather untidy outcome that is still unfolding. The results from the studies by Green and Armstrong suggest that in its current form, game theory has had limited success in forecasting ''how people behave in various circumstances''.

References

Armstrong, J. S. (2002). Is the evidence on forecasting conflicts based on proper science? Working paper.

Binmore, K., & Klemperer, P. (2002). The Biggest Auction Ever: the Sale of the British 3G Telecom Licences, *Economic Journal*, in press.

Bolton, G. (2002). Game theory's role in role-playing. *International Journal of Forecasting*, *18*, 353–358.

Camerer, C. (2002). *Behavioral Game Theory: Experiments on Strategic Interaction*, Princeton: Princeton University Press, in press.

Goodwin, P. (2002). Forecasting games: Can game theory win? *International Journal of Forecasting*, *18*, 369–374.

Green, K. C. (2002). Forecasting decisions in conflict situations: A comparison of game theory, role-playing, and unaided judgement. *International Journal of Forecasting*, *18*, 321–344.

Erev, I., Roth, A., Slonim, R., & & Barron, G. (2002). Predictive value and the usefulness of game theoretic models. *International Journal of Forecasting*, *18*, 359–368.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision making under risk, *Econometrica*, 263–291.

Kaletsky, A. (2002). What the World Needs Now is Economists, *Times of London*, March 28.

McAfee, P., & McMillan, J. (1996). Analyzing the airwaves auction. *Journal of Economic Perspectives, Winter*, *8*, 159–175.

McMillan, J. (1994). Selling spectrum rights. *Journal of Economic Perspectives, Summer*, *8*, 145–162.

Rabin, M. (1994). Incorporating behavioral assumptions into game theory. In Friedman, J. (Ed.), *Problems of Coordination in Economic Activity*. Norwell, MA: Kluwer Academic Publishers.

Raiffa, H. (1982). In *The Art and Science of Negotiation*. Cambridge, MA: Harvard University Press.

Seybold, A. (1995). PCS: The good and the bad news, PEN Computing Magazine, June, 24–26.

Thaler, R. (1988). Anomalies: the ultimatum game. *Journal of Economics Perspectives*, *2*, 195–206.