Embroiled in a conflict: who do you call?

Kesten C. Green*

School of Business and Public Management, Victoria University of Wellington and Managing Director, Decision Research Ltd., P.O. Box 5530, Wellington, New Zealand

Abstract

I address commentators' concerns about the research reported in my paper. These concerns do not threaten the conclusion that role-playing should be preferred ahead of game theory and unaided judgement for forecasting decisions in conflicts. I provide additional evidence and argument that the relative forecasting accuracy of game theory is a legitimate subject for research. I discuss non-forecasting uses for game theory and suggest that, without forecasting validity, such applications may be ill-founded. Replication of the Green research (Green, K. C. (2002) *International Journal of Forecasting 18*, 321–344) by game-theory advocates would be valuable. Extending the research with forecasts for more conflicts would allow greater confidence in recommendations to managers. Extensions should aim to increase the variety of conflicts so that managers can match research findings with their own forecasting problems. More data may allow researchers to identify conditions that favour particular forecasting methods and to estimate the effects of variations in conflict descriptions and decision options. © 2002 International Institute of Forecasters. Published by Elsevier Science B.V. All rights reserved.

Keywords: Conflict; Evaluation; Expert opinion; Forecasting; Game theory; Judgement; Prediction; Role-playing; Simulated interaction; Simulation

I respond to the commentators by answering four broad questions. First, was the research objective achieved? Second, was the objective biased against game theory? Third, what should be the objectives for future research? Fourth, given what we know now, which forecasting approach should a manager who is embroiled in a conflict adopt?

1. Was the research objective achieved?

The purpose of conducting the research was

to determine the usefulness of methods for forecasting decisions in conflicts among small numbers of decision makers. My hope was that the findings would make it possible to offer useful advice to managers on choosing a forecasting method for problems of this type.

1.1. Research design

Armstrong (2002) evaluated my research and concluded that it did well when measured against traditional scientific criteria. Armstrong's evaluation addressed most of the con-

*Tel.: + 64-4-499-2040; fax: + 64-4-499-2080.
E-mail address: kesten.green@vuw.ac.nz (K.C. Green).

cerns of the other commentators over the design and implementation of the research.

After completing the study I reviewed it using the evaluation principles presented in Armstrong (2001b). The study fell short of the ideal for three of the 32 principles. The first of these failings was not to have used objective tests of the assumptions inherent in the conflict descriptions. Truly objective tests for descriptions of human interaction are not possible. Nevertheless, it is good practice to ask people who are involved in a conflict to review and comment on the verity of descriptions and the completeness of decision options, to ask experts to do the same, and to test the material with research participants to ensure they interpret the material in the way that the author intends. Although procedures fell short of the ideal, there is no reason to suppose that any weaknesses in the descriptions or decision options favoured any one forecasting method over any other. A summary of the review is available on the Internet at kestencgreen.com.

Bolton (2002), Erev, Roth, Slonim, and Barron (2002), and Goodwin (2002) took issue with the fact that for four of the six conflict situations the outcome was known to the author of the description. Bolton (2002) rightly pointed out that an author must decide what information to include in the description and implies that an author who knows the outcome may write a materially different description to one that would be written without this knowledge. Although using conflict descriptions written before the outcome was known is desirable, I did not have the resources to obtain role-play and unaided-judgement forecasts for more new situations. Having said this, is there any evidence that remedying this lapse might have affected my conclusions? The answer is 'no'. Both role-play forecasts for the two conflicts that were described by authors *without* knowledge of the outcomes, and role-play forecasts

for the three[1] that were described *with* knowledge of the outcomes, were substantially more accurate than game-theorist forecasts and un-aided-judgement forecasts. In the former case there was a 54 percent error reduction versus game-theorist forecasts and in the latter a 45 percent error reduction. Moreover, the conflict descriptions written by Armstrong and colleagues (Artists Protest, Distribution Channel, and 55% Pay Plan) were not written in the expectation that they would be used to test the accuracy of game theorists' forecasts. There were substantial error reductions between game-theorist and role-play forecasts for these conflicts (44 percent) and for those I had written (56 percent).

The second of the failings was to have conducted only limited testing of the assumptions inherent in the conflict for construct validity. Ideally, one could ask a number of independent domain experts to produce descriptions of a conflict and generate decision options in brainstorming sessions. The forecaster or researcher could then observe the effect of the descriptions on the forecasts obtained. In practice, I did not have the resources to do this and the evidence of Armstrong's (2001a) research and my own is that accurate forecasts can be obtained using one description written by a single independent observer.

The third of the failings was not to have established whether potential clients understand the methods. Ideally, one would survey a diverse sample of managers about this. Armstrong, Brodie, and McIntyre's (1987) survey provided some indirect evidence on the accep-

---

[1]The Panalba situation is excluded from this analysis. As it was described, the situation did not involve interaction between the parties in conflict and hence the Panalba role-plays were not "simulated interactions". See Armstrong (2002) for a discussion on the term and on the exclusion of Panalba from analysis.

tability of various conflict forecasting methods. Research using the Nurses Dispute was supported by a New Zealand government department. Officials found the role-play forecasts sufficiently useful to commission a second study.

1.2. Implementation

Bolton (2002) cautioned against the possible existence of 'lab effects', but did not draw specific conclusions about their existence or otherwise in my research. Armstrong (2002) also addressed this question and found that identifiable biases favour forecasts by game theorists. Other, unidentified, biases might be present, but it is clear from the findings presented in Table 1 in my paper and in the analysis presented in Armstrong's (2002) Table 1, that any unknown bias would have to be very large for its removal to eliminate the advantage in accuracy that the role-play method enjoys.

'Was game theory used?' is a thorny question that elicited a variety of responses from the commentators. As I state in my paper, for my purpose game theory was 'what game theorists do when faced with practical forecasting problems'. My appeal to game-theorists stated that the purpose of the research was to compare game theory with other forecasting methods for conflicts. Why would these experts not have used the skills of their discipline?

2. Was the research objective biased against game theory?

2.1. Game theory is not intended for forecasting real conflicts, or is suitable only for some

The commentators were divided on whether game theory should be regarded as a legitimate method for forecasting real conflicts. Goodwin (2002) found evidence that some game theory researchers regard game theory a legitimate forecasting method. Erev et al. (2002) suggested that there are arenas in which game-theory models produce accurate forecasts. Shefrin (2002) implicitly agreed that game theory can be a legitimate forecasting method when he suggested that game theory's forecasting accuracy would be improved by taking better account of human behaviour. Armstrong (2002)—using the analogy of medical research—made the point that predictive validity is an important test of the value of a theory. To Bolton (2002) and to Wright (2002), on the other hand, testing the relative accuracy of game theorists' forecasts is a little like entering an animal physiologist in the Grand National Steeplechase.

Since writing my paper I have continued to search for evidence on the use and accuracy of game theory for forecasting. Goodwin (2002) found evidence for both on the Decision Insights Incorporated website. I asked Decision Insights founder Bruce Bueno de Mesquita[2] for evidence on the relative accuracy of his expected utility model for forecasting conflicts. He referred me to Stanley Feder of Policy Futures, Frans Stokman of the University of Groningen, and to Stokman and Bueno de Mesquita (1994). I sent e-mail messages to both Feder and Stokman. Stokman did not reply. Feder referred me to Feder (1987) and to the work of Fraser and Hipel (1984) on the use of game theory to forecast conflict outcomes[3]. Feder (1987) compared the accuracy of forecasts using the Bueno de Mesquita model with the accuracy of the

---

unaided judgement forecasts of experts who had provided the input data for the model. He found that both methods were 90 percent accurate.

Stokman and Bueno de Mesquita (1994) compared the accuracy of forecasts of sixteen European Community decisions from the Bueno de Mesquita expected utility model with the accuracy of forecasts from alternative, 'compromise exchange', models. All models used the same variables, but differed in assuming conflict (expected utility model), or exchange (alternative models). The Bueno de Mesquita model's forecasts were more accurate on average than those from the other models on the basis of the measures that were used by the authors. The differences in accuracy between the Bueno de Mesquita model forecasts and the forecasts from the best of the alternative models were not great, however. More importantly for the objective of my research, the authors do not compare the accuracy of the model forecasts with the accuracy of forecasts from experts using unaided judgement—the approach that is typically adopted for such problems.

I inspected Fraser's publication list[4] and found one title which included the word 'forecast': Fraser (1986). Although both this work and Fraser and Hipel (1984) promote the game-theoretic technique 'conflict analysis' as a forecasting method for any conflict situation, neither offered evidence on comparative forecasting accuracy. Fraser did not respond to my email message of 5 December.

While opinions on the use of game theory as a forecasting method are diverse, some researchers do believe this to be a legitimate subject for research. Moreover, practitioners use game theory for forecasting. The relative performance of game theory for forecasting conflicts is clearly an important issue.

---

[4]At http://www.openoptions.com/publications.htm on 5 December 2001.

2.2. Game theory is useful for describing situations

Some of the commentators discuss the value of game theory for applications other than forecasting, implying that game theory can be useful irrespective of its forecasting accuracy. It was not the objective of my research to investigate the non-forecasting uses of game theory. Nevertheless, the use of game theory to describe situations (Bolton, 2002) and to design markets (Erev et al., 2002; Shefrin, 2002) warrants responses.

Bolton (2002) suggested that role-play forecasters need the guidance of game theory in writing descriptions. Game theory was not used in designing the material for any of the six conflicts used in my research and yet an absolute majority of role-play decisions matched the actual decision for five of the six conflicts. It is nevertheless possible that the use of game theory for this purpose may lead to an increase in the accuracy of forecasts.

2.3. Game theory is useful for designing markets

Erev et al. (2002) and Shefrin (2002) discussed the value of game theory for designing markets. In a mechanical world, this would not involve forecasting: the designer would set the rules confident that the range of possible behaviour and outcomes were circumscribed.

The spectrum auctions in the United States have been seen as a showcase for game theory. It was used by market designers and participants, and billions of dollars were at stake. Shefrin's (2002) description of the auctions and their aftermath suggests that the outcomes were mixed, at best. There was more money raised than had been anticipated. On the face of it, this was a bonus for tax-payers. But tax-payers are also investors, employees, and users of the carriers' services. Outcomes mentioned by Shef-

rin (2002) include massive losses and plummeting stock prices among carriers, defaults by and bankruptcy of successful bidders, layoffs, and ongoing litigation and negotiation. Perhaps the spectrum could have been allocated in better ways. Why were the undesirable outcomes not anticipated by the game-theorist consultants?

3. What should be the objectives for future research?

3.1. Replication

As Armstrong (2002) suggested, replication by other researchers, particularly by game-theory experts, would be valuable.

3.2. Classification of conflicts, and extension

Armstrong (2002) suggested studying the conditions under which each forecasting method is most accurate. Both Erev et al. (2002) and Goodwin (2002) discussed the possible existence of a subset of conflicts that are particularly amenable to game-theoretic forecasting.

Erev et al. (2002) proposed using a classification scheme from which to draw representative samples of conflicts with which to compare the forecasting accuracy of methods. This is an interesting idea, and it is good practice to use a variety of conflicts to test the methods (Principle 13.14, Armstrong, 2001c), but what does 'representative' mean in the context of real conflicts? Classification criteria are, in this context, arbitrary. The Erev et al. (2002) criteria are also abstract, and the authors do not attempt to match the six conflicts used in my research with six from their own classification scheme. Shefrin (2002) did attempt matching, drawing parallels between the Dollar Auction game from the game-theory literature and the two conflicts 55% Pay Plan and Nurses Dispute. He concluded that the 'traditional' game-theoretic ap-

proach will not produce accurate forecasts of these conflicts.

An alternative approach to that of Erev et al. (2002) would be to compare the accuracy of forecasting methods using conflicts belonging to categories that are understood by, and relevant to, managers. For example, conflicts that are diverse in arena (ie, industrial, commercial, civil, political, diplomatic, military) and parties (ie, individuals, organisations, governments). Categories such as these would allow managers and practitioners to match their forecasting problems with research findings.

3.3. Conflict description

Will conflict descriptions written by independent authors tend to lead to the same forecast? A formal framework, such as that proposed by Bolton (2002), may help to improve the validity of descriptions.

Does the framing of decision options affect forecasts? Goodwin (2002) raised concerns about the effect of decision options on strategy formulation. Possible decisions can be identified by, for example, asking protagonists and experts, and by conducting brainstorming sessions. In practice, role-players, or those using other methods to forecast, may propose alternative decisions. Experts or protagonists could judge the viability of any such forecast decisions. These are all matters for further research.

Does the number of decision options affect forecast accuracy? There currently is no clear evidence on this question.

3.4. Conflict experts' forecasts

The unaided-judgement forecasts described in Green (2002) were largely obtained from naive judges (mostly students) whereas game theory forecasts were obtained from experts. Armstrong (2002) suggested asking experts on

conflicts or on forecasting, who are not game-theory experts, to provide forecasts.

Wright (2002) cited evidence that expert judgement can only be expected to be of more value than the judgement of novices when the subject of judgement has both *ecological validity* for the expert and *learning* is possible. He suggested that neither requirement holds for the game-theory experts and postulates this as an explanation for their performance. The first requirement (ecological validity) is equivalent to asking 'is this person an expert *in this field*'. Game theorists are experts in the study of conflicts. It seems reasonable to assume that the conflicts used in the research had ecological validity for them. In fact, the game-theorist participants were unique in that they volunteered their services, as experts, for forecasting these conflicts. The expertise of someone who was familiar with New Zealand industrial relations might have been more relevant for forecasting the Nurses Dispute, but this is not the same as saying that the expertise of a game theorist who is not familiar with New Zealand industrial relations is *ir*relevant to this problem.

Wright (2002) suggested that his second requirement (the possibility of learning) was not met because the novelty of the conflicts precludes learning being brought to bear on the forecasting problems. If he is right in this, the role of game theorists, and of other conflict experts, is limited.

Each role-play forecast requires several participants, whereas forecasts from the other methods were mostly obtained from individuals working in isolation. Collaboration requires forecasters to justify their forecasts to their fellows and allows forecasts to be combined. Both justification and combining tend to increase the accuracy of judgmental forecasts (Stewart, 2001). Obtaining collaborative and individual unaided-judgement and game-theory forecasts would allow the effect of collaboration on conflict forecasting accuracy to be estimated for these methods.

3.5. Understanding role-play

While other commentators were chiefly concerned with establishing how or whether game theory can better be applied to forecasting conflicts, Armstrong (2002) and Wright (2002) both argued for expending more research effort on role-playing.

Although, in contrast to Wright (2002), Shefrin (2002) suggested that the quirks of human behaviour *can* be incorporated into game theory (game theorists can learn), he speculated that role-play forecasts will tend to be more accurate because the method uses actual interactions between real people subject to emotion and errors of judgement.

4. So, who *do* you call?

Research needs to be done on forecasting decisions in conflict situations. Nevertheless, compelling evidence exists that role-playing will provide more accurate forecasts than other methods, and the cost is modest. Role-playing is a successful and efficient way for forecasters to take real human behaviour into account.

References

Armstrong, J. S. (2001a). Role playing: a method to forecast decisions. In Armstrong, J. S. (Ed.), *Principles of forecasting: a handbook for researchers and practitioners*. Norwell, MA: Kluwer Academic Publishers, pp. 15–30.

Armstrong, J. S. (2001b). Evaluating forecasting methods. In Armstrong, J. S. (Ed.), *Principles of forecasting: a handbook for researchers and practitioners*. Norwell, MA: Kluwer Academic Publishers, pp. 443–472.

Armstrong, J. S. (2001c). Standards and practices for forecasting. In Armstrong, J. S. (Ed.), *Principles of forecasting: a handbook for researchers and practitioners*. Norwell, MA: Kluwer Academic Publishers, pp. 679–732.

Armstrong, J. S. (2002). Assessing game theory, role playing, and unaided judgement. *International Journal of Forecasting*, *18*, 345–352.

Armstrong, J. S., Brodie, R. J., & McIntyre, S. H. (1987). Forecasting methods for marketing: review of empirical research. *International Journal of Forecasting*, *3*, 335–376.

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

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

Feder, S. A. (1987). Factions and Policon: new ways to analyze politics. *Studies in Intelligence*, *31*(1), 41–57.

Fraser, N. M. (1986). Political and social forecasting using conflict analysis. *European Journal of Political Economy*, *2*(2), 203–222.

Fraser, N. M., & Hipel, K. W. (1984). *Conflict analysis: models and resolutions*. New York: North-Holland.

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.

Shefrin, H. (2002). Behavioral decision making, forecasting, game theory, and role play. *International Journal of Forecasting*, *18*, 375–382.

Stewart, T. R. (2001). Improving reliability of judgmental forecasts. In Armstrong, J. S. (Ed.), *Principles of forecasting: a handbook for researchers and practitioners*. Norwell, MA: Kluwer Academic Publishers, pp. 81–106.

Stokman, F. N., & Bueno de Mesquita, B. (1994). Models of exchange and of expected utility maximisation: a comparison of accuracy. In Bueno de Mesquita, B., & Stokman, F. N. (Eds.), *European community decision making: models, applications, and comparisons*. New Haven, CN: Yale University Press, pp. 214–228.

Wright, G. (2002). Game theory, game theorists, university students, role playing and forecasting ability. *International Journal of Forecasting*, *18*, 383–387.