Preliminary
Comments welcome

Extremists' Beliefs and Beliefs about Extremists: Common Knowledge in the Wake of Acts of Mass Terror*

Sanjay Jain†
Department of Economics,
University of Virginia,
114 Rouss Hall,
Charlottesville, VA 22903, U.S.A.
Tel: +1 (434) 924-6753
Fax: +1 (434) 982-2904
e-mail: sjain@virginia.edu

Sharun Mukand
Department of Economics,
Tufts University,
305 Braker Hall,
Medford MA 02155, U.S.A.
Tel: +1 (617) 627-5476
e-mail: smukand@tufts.edu

June 2002

version 2.1

* Paper prepared for the workshop on "The Economic Consequences of Global Terrorism", 14-15 June, DIW Berlin, Germany. For helpful comments and discussions, we are grateful to Simon Anderson, Albert Choi, Maxim Engers, Sumit Joshi, and Sandeep Kapur. However, all views expressed in this paper, and any remaining errors, are those of the authors alone, and should not be ascribed to anyone else.
† Corresponding author.

Abstract

In this note, we analyze two distinct but related questions, from a game-theoretic viewpoint. First, we analyze some implications of very visible, discrete, large-scale terrorist actions, such as the September 11 attacks. The key element of such incidents is that they become part of the set of common knowledge, in ways that lesser incidents do not. Thus, for example, changed beliefs about the likely motives of future airplane hijackers may result in security forces, flight staff, passengers, etc, coordinating around a strategy of 'fight' (in the hope of stemming greater losses) rather than to 'negotiate'. In turn, this change in the likely actions of the other parties will change the composition of the pool of hijackers, shifting it towards the most extreme terrorists, which in turn supports the equilibrium beliefs of all the parties.

A second, related, issue that we examine also concerns the appropriate policy response to acts of terrorism. Should governments negotiate with terrorists, or insurgent movements more broadly? What sort of stance should they take? As a consequence of the time consistency problems inherent in government promises not to negotiate with terrorists, governments often delegate the decision on whether to negotiate or not, and the negotiating stance to take. Our analysis highlights one trade-off that governments face in choosing this 'delegate'. On the one hand, choosing a 'tough' negotiator *ex ante* may yield a better bargaining outcome *ex post*. On the other hand, the choice of a tough, unyielding delegate might also have a signaling effect, so that the insurgents might conclude that the initiation of the negotiations is merely a diversionary tactic while the government prepares itself to fight. In that case, the likelihood of a compromise in any negotiation is reduced, for a tough, uncompromising delegate, which in turn might moderate the government's choice of delegate. This effect, to the best of our knowledge, has not been systematically analyzed in the literature, especially in the context of hostage-taking and terrorism.

Keywords: common knowledge, co-ordination, strategic delegation, negotiation, terrorism.

I. Introduction

In this note, we analyze two distinct but related questions, from a game-theoretic viewpoint. First, we examine some implications of very visible, discrete, large-scale terrorist actions, such as the September 11 attacks. The key element of such incidents is that they become part of the set of common knowledge, in ways that lesser incidents do not. Using a game-theoretic model, we analyze the impact on the beliefs and actions of the affected parties, such as security forces, airline employees and passengers, etc, as well as on the beliefs and actions of terrorists, along with the resultant change in the pool from which they are drawn. For instance, changed beliefs about the likely motives of future airplane hijackers may result in security forces, flight staff, passengers, etc, coordinating around a strategy of 'fight' (in the hope of stemming greater losses) rather than to 'negotiate'. (For example, prior to September 2001, standard operating instructions to commercial airline pilots were to comply with hijackers' instructions until contact with negotiators could be established. These instructions have since been changed). In turn, this change in the likely actions of the other parties will change the composition of the pool of hijackers, shifting it towards the most extreme terrorists, which in turn supports the equilibrium beliefs of all the parties.

Further, we suggest that our analysis of this phenomenon of the 'erosion of the middle', and the shift towards more extreme actions, sustained by self-fulfilling beliefs, can be applied more generally. For example, the taking up of arms by an insurgent or dissident movement, even those that do so solely as a means of pressuring governments to the negotiating table, is likely to be met by the much greater use of force by governments. Anticipating this, only those movements that, in any case, have a greater inclination toward violence are likely to adopt it as a strategy, thus justifying the government's beliefs in equilibrium.

The second, related, issue that we examine also concerns the appropriate policy response to acts of terrorism. Should governments negotiate with terrorists? What sort of stance should they take? As Sandler and Enders (2002) point out, a refusal to negotiate with terrorists is one of the cornerstones of most countries' policy. However, given the obvious time inconsistency problems that this policy creates, governments often 'capitulate' (to use the terminology of Lapan and Sandler (1988)), with the associated reputational and other consequences. A more realistic analysis should incorporate this possibility, and the various channels through which governments might reduce the likelihood of this happening, for example, by investing in deterrence efforts.

We offer a complementary argument. As we suggest below, in a more general context, there is an additional element that might be important in the government's decision-making on negotiations. This is the choice of the negotiator, or more broadly, the 'delegate' to whom the government leaves the decision on whether to negotiate or not, and what stance to take in any negotiations. Our analysis highlights one trade-off that governments face in choosing this delegate. On the one hand, choosing a 'tough' negotiator *ex ante* may yield a better bargaining outcome *ex post*. (This effect is well-known in the literature on strategic delegation, especially in the context of bargaining

between a labor union and management). On the other hand, the choice of the delegate might also have a signaling effect, so that the likelihood of a compromise in any negotiation is reduced, for a tough, uncompromising delegate. This might lead the government to be moderate its selection, so that it might choose a less 'extreme' delegate. This effect, to the best of our knowledge, has not been systematically analyzed, especially in the context of hostage-taking and terrorism.

The paper is organized as follows. In the next section, section II, we briefly discuss the relevant related literature, and place our insights relative to the results available in that literature. In section III, we develop our analysis of the impact of acts of mass terror. In section IV, we consider the implications of strategic delegation in negotiating with terrorists. Section V concludes, with some suggestions for future research.

In what follows, our emphasis is on providing illustrations of the way that existing and well-understood concepts can be applied to the analysis of a new set of issues, having to do with global terrorism, in often unexpected ways. Our objective in this paper is not the creation of new formal models per se, but rather the development of insights by adapting existing theory. Accordingly, our analysis is somewhat informal – since the underlying formal theory is well-developed, and quite well understood, we concentrate instead on the insights that can be gleaned from drawing on existing theoretical tools and applying them to new contexts. In choosing the simplest possible models to convey out arguments, we are consciously sacrificing generality for simplicity, and what we hope is a clarity of exposition.

II. Related literature

The related literature in this area, on the application of game-theoretic tools to model decision-making in the context of hijackings and hostage-taking, is remarkably thin. (For a survey, see Sandler and Enders (2002)). This is especially surprising given the obvious importance of formulating policy responses, to counteract and respond to terrorism, even before the September 11 attacks, and the obvious applicability of available tools in the theoretical literature on game theory and models of asymmetric information. For example, the survey by Sandler and Enders (2002) identifies a mere handful of papers that deal with the economic analysis of hostage-taking events, of which only three utilize game-theoretic tools in any depth. Of these, this note is perhaps closest in spirit to the analysis in Lapan and Sandler (1988), which analyzes the credibility of government commitments to never negotiate with terrorists, and makes the point that such promises may be time-inconsistent, even when there are reputational costs to reneging on one's commitments. In some sense, our approach is complementary to the approach taken by Lapan and Sandler (1988). For one thing, they allow the government to choose its level of deterrence expenditure, which affects (negatively) the likelihood of logistical success by the terrorists. This is not an issue we consider, focusing instead on the effect that the choice of negotiator has on the final outcome. By allowing the government to choose a negotiator, we are in a sense endogenizing the decision on whether to 'capitulate' or not.

The importance of having governments co-ordinate their response to terrorist actions is emphasized in a couple of lines of research. For example, in choosing the levels of deterrence effort (or expenditure) to undertake, there is a certain 'transference externality' that is imposed by countries on each other when countries choose those deterrence levels independently of each other. This externality could be positive or negative, depending on the relative strength of two opposing effects: on the one hand, country A's deterrence efforts make country B's citizens safer (by making attacks on them less likely in country A). On the other hand, increased deterrence in country A also shifts the attacks abroad, to country B. The analysis of remedies again emphasizes the importance of increasing co-ordination among countries, and points out that a piecemeal approach is likely to be unsuccessful. (For a discussion, see Sandler and Enders (2002)).

Similarly, in choosing retaliatory action against the perpetrators of terrorism, co-ordination may be key. Typically, the cost of retaliation (in terms of attracting subsequent terrorist actions) is privately borne, whereas the benefit accrues to all countries. In this situation, the classic prisoners' dilemma prediction would be that all countries would under-provide a resource that has value to all. (Lee (1988), Sandler and Enders (2002)). Again, as with most public good provision problems, a coordinated response can go some way to solving this problem.

Our analysis of the co-ordination problem differs from these analyses in that we focus on implicit rather than explicit co-ordination. In other words, there is no overt co-ordination among the concerned parties, for example, among passengers and flight crew choosing whether to attack aircraft hijackers. Such communication may simply be impossible in the circumstances. Thus, convergence to a (new) equilibrium happens because the parties independently update their beliefs, and take actions accordingly, so that these beliefs turn out to be self-fulfilling in equilibrium. And all this happens in the absence of any formal co-ordinating agency.

In the next section, we look at the implications of changes in common beliefs on the likely pattern of terrorist actions.

III. Common Knowledge

Consider the following game. In stage 1, a potential terrorist decides whether or not to hijack an aircraft in order to achieve his objectives. Suppose potential hijackers come in two types: one, whom we dub 'fanatics', intend to use the aircraft on a suicide mission, akin to the September 11 hijackers. The second type are the classic 'hostage-takers', whose intention is to use the aircraft and its onboard crew and passengers as bargaining chips in order to win some political or criminal objective (for example, the release of their previously captured colleagues from prison). The obvious distinction is that negotiating with the first kind of hijacker is futile, whereas the very objective of the second type is to get to a stage where he can negotiate for his demands to be satisfied. Further, we make the reasonable assumption that these two types are visually indistinguishable in the sense that if the fanatic asserts that he is a hostage-taker, his true

type cannot be known with certainty, until he chooses to reveal it by his actions (for example, by trying to crash the aircraft).

We now consider the following scenario. Suppose a potential hijacker (or a group of hijackers) manages to successfully hijack an aircraft. How should the affected parties respond? These parties include the passengers, the flight crew, the airline, the airport authorities, law enforcement authorities, and more generally the government. For simplicity, we examine only the immediately affected parties, the people on board the plane. For further simplicity, we will refer to them as 'passengers', although of course they could also include the flight crew, and any law enforcement officers (such as the US 'Sky Marshals'). A final simplification is to model this as a two-passenger game, and analyze the co-ordination problem that they face.

Suppose that the passengers can adopt one of two actions: co-operate, or resist.[1] The payoff matrices corresponding to these actions are given in Tables 1. The interpretation is straightforward. There are two possible games that the passengers could be involved in. In game F, the 'state of nature' is that the hijacker's true type is 'fanatic'. In game H, the hijacker's true type is 'hostage-taker'. The true state of nature is unknown to the passengers. In both states, the passengers do best by co-ordinating their actions. However, the 'correct' action to co-ordinate on, differs depending on which game is being played. It is dangerous for a passenger to resist (i.e., attack the hijacker) if his counterpart is not also resisting.[2] The payoffs can be viewed as the expected values associated with a particular set of actions. So, for example, the payoffs from (resist, resist) can be thought of as the weighted average of the payoff in the event that the resistance to the hijackers is successful, and if it is unsuccessful, respectively, with exogenously given known probabilities associated with the respective events.

Suppose that initially, the game H is the more likely one, and the state F is far more unlikely, with probability p that is very small. In that case, if the players can not communicate, then they can do no better than to play (co-operate, co-operate), and get an expected payoff of (1-p).M – p.D each (assuming that D is not overwhelmingly bigger than M, so that the expected payoff is greater than 0 each).[3]

The question of interest to us is: what implications do acts of mass terror, such as those of September 11, have for this co-ordination game? The key distinction between such widely-publicized, major events and other less cataclysmic attacks is that the former can reasonably be supposed to be common knowledge among the passengers and the hijackers, while the latter are not. (For brevity, we shall refer to the former kinds of major events as acts of 'mass terror', as distinct from 'isolated attacks'). This distinction is important: as Geanakoplos (1992) points out, and Rubinstein (1989) demonstrates, the

---

[1] In what follows, we speak only in terms of pure strategies. We speculate that the analysis would be no different, at least qualitatively, if we were to also consider mixed strategies.

[2] Let D>L >M. The reader might recognize that the basic payoff structure is similar to the coordinated attack problem in Rubinstein's (1989) classic paper. For an exposition, see Osborne and Rubinstein (1994). However, the timing and information structure described here is very different.

[3] This does not seem unreasonable, if one interprets, somewhat crudely, -D as the disutility associated with death, and M as the utility associated with continued life.

gulf between 'common knowledge' and 'almost common knowledge' is a very wide one. In our game above, passenger A might know that some other incident has occurred (e.g., that in an 'isolated attack', another aircraft was hijacked by fanatics at some time in the past), and has adjusted his beliefs (say, from p to q) about the hijacker's true type. Similarly, passenger B might know of the same incident and also know that passenger A knows of it, but does not know if passenger A knows that B knows about that previous incident. In this case, the incident is not common knowledge, and the associated updating of beliefs about the hijacker's true type is also not common knowledge. But an attack of the magnitude of the September 11 attacks can legitimately be assumed to be part of the set of common knowledge.

One implication of this is straightforward. If each passenger updates his beliefs about the probabilities associated with hijacker types from p to q after an even of mass terror that is common knowledge, then the optimal actions in the co-ordination games above change. The expected payoff from (co-operate, co-operate) now declines, so that now (resist, resist) may be optimal, i.e., $(1-q).M - q.D > 0$, where $q > p$. Is the updating of beliefs from p to q rational? To some extent, any updating of beliefs when new evidence comes in that an event has occurred (especially a low probability one) is sensible. But a more interesting insight may be gained by asking: what impact does this updating of beliefs have on the potential hijackers? Post September 11, it seems reasonable to argue that passengers are likely to assign a higher probability to the state that the hijackers are fanatics, and are far more likely to resist rather than co-operate. Suppose that (resist, resist) is now the optimal action by passengers. In that case, even without specifying the payoff functions of the two different types of potential hijackers, it is clear that the hijackers can take as given that the denouement of the hijacking episode will involve violence. Assuming that the propensity for violence, relative to negotiating for ransoms or release, is higher among the fanatics than among the hostage-takers, then a successful act of mass terror by the fanatics ends up conferring a significant negative externality on the hostage-takers. In some sense, by shifting the passengers' actions to (resist, resist), the fanatics have raised the costs of hijacking to the hostage-takers.

Thus, the pool of hijackers itself will change, as the hostage-takers 'exit' this particular activity and turn to other strategies to press their claims. This then has the consequence that the updating of beliefs by the passengers was perfectly sensible for another reason: viz., that the hijacker pool has in fact shifted towards the fanatics.

Isolated attacks do not have the same impact as an act of mass terror, for the reasons discussed above. The pool of hijackers changes because passengers co-ordinate on the resistance actions, but that co-ordination is made possible because acts of mass terror, and the associated updating of beliefs, is common knowledge in a way that isolated attacks are not.

There are some counterintuitive predictions that emerge from this discussion. For one, if the hostage-takers 'exit the market', then the number of hijackings should drop. Note that this effect, reducing the incidence of hijackings, is independent of the anticipated reduction in hijackings because of increased security measures at the airports. In other

words, even in the absence of those increased security measures, one should expect hijacking incidents to decline. Somewhat paradoxically, the attacks of September 11 may actually have decreased the likelihood of being hijacked!

If the incidence of hijackings declines, then the hostage-takers will probably adopt other methods to press their claims – perhaps even by seizing hostages in some other manner. In formulating policy responses to terrorism, as Sandler and Enders (2002) point out, it is important to recognize what they refer to as these 'substitution possibilities'. The discussion above supports their argument that antiterrorist policies must be pursued on a wide front, so as to get terrorists to adopt less harmful strategies, by raising the cost of terrorism across the board, rather than only for some kinds of terrorism.

IV. Negotiation via a delegate

So far, we have considered individual, decentralized, responses to terrorism. In this section, we examine the determinants of the stance that governments should take in dealing with terrorists. In recognition of the time-inconsistency problem that plagues government promises to 'never negotiate with terrorists' (Sandler and Enders (2002)), governments can and do, in effect, tie their own hands by delegating this decision to an independent agency or authority. Obviously, for this commitment to be meaningful, the delegate has to have the authority to make commitments, or sign 'deals', on the government's behalf. In our analysis, we assume that the delegate has this authority, so that in choosing the delegate, the government knows that their agent's preferences might be different from that of his principal, the government.

We begin by asking whether governments (or societies, more generally) will necessarily choose a delegate whose preferences closely reflect their own. We argue that, for strategic reasons, the principal may well choose an agent whose preferences are different from their own. The intuition behind this result can be traced back at least to Fershtman and Judd (1987), who showed that a principal (a firm owner) could choose a contract for the agent (the firm manager) that rewarded the agent for maximizing a different variable than the one that the principal truly cares about (sales versus profits). Essentially, in choosing the contract type for the agent, the principal can affect the competition game that is played in the second stage between his agent and other firms.

We construct a simple set-up to develop a similar intuition in a negotiation framework. Consider a two stage game, in which the government announces its choice of delegate (or authorizes a particular government agency) for all forthcoming negotiations involving, for example, insurgent movements. In the second stage, the delegate negotiates with the other party, say the insurgents. Let the 'negotiating space' be represented by a line segment from 0 to 1. Suppose that the government's optimal position is at the mid-point value of ½. Further, suppose that the insurgents, the counterpart in the negotiations to come, are located at 0. The government can, however, choose a delegate to negotiate on its behalf. Each potential delegate's favored point is public knowledge, and we let d denote the delegate's location on the line. Once the government has chosen a delegate,

the second stage negotiation is extremely simple – we use a slightly modified Nash bargaining solution to derive the outcome of the second period negotiation.

Assume that both parties get zero utility in the event that no agreement is reached. Let the utility of each party involved (government, delegate, and the insurgents) be decreasing linearly in the distance between their ideal point and the negotiation outcome, X*(d). So, for example, the delegate seeks to maximize his own utility, $U_d(X^*) = 1 - (d - X^*(d))$. Note that the second term should, strictly speaking, be the absolute value of $d - X^*(d)$, but since for all practical purposes, X*(d) will always be less than d, we write the expression in the more intuitive form above. Similarly, let the insurgents' utility be given by $U_i(X^*) = 1 - X^*(d)$. It is straightforward to observe that the Nash bargaining solution in this situation would deliver $X^*(d) = d/2$. We modify the solution as follows: we assume that there is an asymmetry in bargaining power between the delegate and the insurgents, so that the modified solution is halfway between the Nash solution and the ideal point of the delegate, i.e., we let $X^*(d) = d.3/4$. This is a somewhat ad hoc way of capturing the asymmetry, but it has the virtue that it is simple to interpret and analyze. We are now in a position to calculate d*, the delegate that the government will choose. This decision is straightforward, once we note that the final outcome, X*(d) is (3/4).d. In order to achieve $X^*(d^*) = ½$, the government chooses $d^* = (2/3).d$.

We summarize this in our first observation.

Observation IV.1: *The government chooses a delegate whose preferences are different from its own. The delegate is more 'extreme' than the government, in the sense that he is further away from the other party, the insurgents.*

Hence, the government may choose a 'hawk' to negotiate on its behalf, because the hawk can deliver better outcomes in the bargaining stage than a more 'dovish' delegate whose preferences more closely reflect the government's. Note that there is a moderating force at work, too, so that the government does not choose the most extreme hawk, d = 1. While that extremist can deliver outcomes that are even further away from the insurgents, that outcome is too extreme for the government. Hence the government will 'rein in its Rambos', in a manner of speaking.

But there may be an additional moderating influence on the government's choice of delegate. While a more hawkish delegate can deliver outcomes closer to the higher end of the line segment, intuitively it seems reasonable to argue that the chances of the negotiations breaking down without a compromise agreement are higher the more hawkish is the delegate. In that case, the government's expected utility from its delegate can be represented as $EU_g(d) = P(d).(X^*(d) - ½)$, where P(d) is the probability that agreement is reached, with $P'(d) < 0$.

Why might P(d) be decreasing in d? Here, we provide an intuitive description of what such a model might look like. Suppose that, if it were known in advance that no compromise would be reached, then the insurgents would strictly prefer to 'fight' rather than negotiate. Further, suppose that while the negotiations are ongoing, the

government's becomes 'stronger' for the subsequent fight, if the negotiations were to end without a compromise, and such a fight were to take place. In that case, the identity of the delegate might convey useful information to the insurgents, who might attempt to gauge whether the government is serious about negotiating, or whether the initiation of the negotiations is merely a diversionary tactic while the government prepares itself to fight. These leads to our second observation.

Observation IV.2: *The government chooses a delegate whose preferences are different from its own. But the difference may be reduced if the likelihood of success in the bargaining is lower for more extreme delegates.*

The choice of a tough delegate by the government might lead the other party, the 'insurgents', to switch from 'negotiate' to 'fight', thereby dooming the negotiations, which in turn might moderate the government's choice of delegate. Note that the intuition so far has not relied on attitudes towards risk at all. If, in addition, one were to think of the government as being risk-averse, then it may be even less willing to chance the breakdown of negotiations, and moderate its selection of delegate further. In some sense, the more risk averse the government, the more willing it is to give up the strategic advantage that is conferred on it by its ability to delegate.

V. Conclusions

In the preceding sections, we have argued that some events have a 'polarizing' effect, so that in a situation with potential multiple equilibria, such an event can cause a shift from one equilibrium to another. Further, we speculate that our analysis of this phenomenon of the 'erosion of the middle', and the shift towards more extreme actions, sustained by self-fulfilling beliefs, can be applied more generally. For example, the taking up of arms by an insurgent or dissident movement, even those that do so solely as a means of pressuring governments to the negotiating table, is likely to be met by a much greater use of force by governments. Anticipating this, only those movements that, in any case, have a greater inclination toward violence are likely to adopt it as a strategy, thus justifying the government's beliefs in equilibrium.

Other directions for future research include the application of bargaining models to the analysis of the problems discussed here. A rich body of literature on strategic delegation offers insights on the circumstances under which a principal can enhance outcomes by publicly delegating some negotiating authority to an agent, whose preferences (or the objective function in the agent's contract) might be different from his own (Fershtman and Judd (1987)). These modes of analysis should be fruitful in examining the issues discussed in this paper.

References

Fershtman, Chaim and Kenneth L. Judd, 1987, "Equilibrium incentives in oligopoly", *American Economic Review*, vol. 77, pp. 927-940.

Geanakoplos, John, 1992, "Common knowledge", Journal of Economic Perspectives, vol. 6, no. 4, pp. 53-82.

Lapan, Harvey E. and Todd Sandler, 1988, "To bargain or not to bargain: that is the question", *American Economic Review Papers and Proceedings*, May 1988.

Lee, Dwight R., 1988, "Free riding and paid riding in te fight against terrorism", *American Economic Review Papers and Proceedings*, May 1988.

Osborne, Martin and Ariel Rubinstein, 1994, A Course in Game Theory, MIT Press, Cambridge, MA.

Rubinstein, Ariel, 1989, "The electronic mail game: strategic behavior under "almost common knowledge", *American Economic Review*, vol. 79, no. 3.

Sandler, Todd and Walter Enders, 2002, "An economic perspective on transnational terrorism", mimeo.

Table 1.

Game H

	Passenger A		
Passenger B		Co-operate	Resist
	Co-operate	M, M	0, -L
	Resist	-L, 0	0, 0

Game F

	Passenger A		
Passenger B		Co-operate	Resist
	Co-operate	-D, -D	-L, -L
	Resist	-L, -L	0, 0

	Passenger A Co-operate	Resist
Hold	0, 0	0, -1
Sell	-1, 0	3/2, 3/2

(A) Trader Resources = (6, 6) and Govt. Reserves R = 10

	Trader B Hold	Sell
Trader A Hold	0, 0	0, 4
Sell	-1, 0	7/8, 17/8

(B) Trader Resources = (6, 10) and Govt. Reserves R = 10