

Robert Jervis – “Rational Deterrence: Theory and Evidence”

Among the most important questions in social science are the causes and effects of threats and force. For social scientists seeking to study this, the next great question is methodological, discussing the relative merits of deduction and induction for theory building, and the merits of assuming that actors can be seen as seeking to maximize their subjective expected utilities (SEU).

It is important to note that case studies are required to test any model, and that SEU requires a good deal of input about the situation the actor thinks he is facing, what options he perceives, how he ranks his goals, and how he thinks others will react.

The choice between the deductive and approach and one building upon case studies involves a trade-off between “rigor and richness.” The former misses many nuances of individual cases, while the latter will fall short of generalization.

It also needs to be noted that: deductive theories need not imply rationality, and SEU can be used without generating new theories; expected utility does not ignore individual level analysis; methodological approaches cannot be associated with only one set of substantive arguments.

Rational deterrence theory (RDT) is in many ways: it is seen as both prescriptive and descriptive, it is said to deal with national behavior and international outcomes. The ambiguities and the fact that there is not one, but many RDTs makes it very hard to disconfirm RDT.

The best defined RDT, the “classical” or “second-wave” form, includes several assumptions, including: that the state’s adversary is motivated to expand, that it knows the state is defensive, and that leaders on both sides are primarily concerned with the external situation. Much evidence suggests that the last of these is often untrue. There is also evidence that aggression can be motivated by fear of loss as well as desire for gain and, while this has been an insight of RDT, use of second-wave deterrence theory would not always maximize utility. These findings could be included in a broader RDT, but not the second-wave variety. The primary concern of these theorists was that status quo states would lack the strength, skill or resolve to resist, or that the aggressive power would mistakenly believe that they lacked one of these qualities. They did not look at how threats might set off a spiral of increasing threat and conflict rather than deter a would-be aggressor. While a more sophisticated RDT could include this, it would not be particularly parsimonious nor could it be concerned predominantly with threats, force, and coercion.

In looking at deterrence, most social scientists find examples of the outcomes of interest and look backwards from them, and do not look at cases where other outcomes appear. This makes an assessment of actual effects of variables uncertain. Thus in looking at examples of successful and unsuccessful deterrence, the starting point is often after general deterrence has failed to prevent a crisis. In looking at general deterrence, one runs into problems of judging adequate time frame, although often frequency is not that important as looking at the actual cases of failure, as the failure of deterrence in a case when there were proper conditions for it would be highly damaging, while the question of why statesmen undertake dangerous courses may be

particularly interesting.

Generalizations from case studies are limited. There are many idiosyncracies that cannot be easily explained by generalities. Still a large number can be drawn from three broad categories:

1. Cognitive limits on information processing, including use of existing theories and analogies, as well as tendencies to underestimate confusion and make simplified decisions.
2. Motivated biases, stemming from psychological pressures and needs, including the tendency to think well of ones state, and the tendency to view situations as more promising than they are.
3. The role of domestic politics and organizational behavior in affecting how policies are formed and implemented.

These findings have important implications on deterrence:

1. Adversaries seldom understand each others goals, fears, means-ends beliefs, and perceptions.
2. The adversary will miss or misperceive many of the state's signals.
3. Even commitments by one actor that are objectively clear and credible may not be perceived.
4. Actors tend to overestimate the potency of threats and underestimate the utility of rewards and reassurances.
5. Threats and conciliation must be combined, but the optimal mixture is hard to gauge.
6. The success of immediate deterrence may weaken future general deterrence, by making the situation more undesirable for the adversary or increasing the perception of the states hostility.

Key question is whether an RDT (albeit one not yet developed) could explain the findings from case study research. Here it is critical to see if SEU takes actors' beliefs as givens, or asks if they are reasonable.

Even an SEU considering information-processing costs would expect that in important circumstances, cognitive biases would be limited, but this has not been demonstrated. An SEU might be created to include calculations of psychological pain caused by accepting hard facts and assuming a high (and constant) discount rate for future benefits. This would be a rather unsatisfactory theory that would seem to stray rather far from any extant RDT.

The minimum requirement of SEU is that behavior and thus inferred preferences be consistent. This is not the case, as coalitions are often built and maintained through the use of contradictory policy measure; disagreements can be so deep that the process is deadlocked; inconsistencies will appear as different factions come to power; divergent bureaucratic interests and perspectives make formation and (even more so) implementation of policy inconsistent. In fact, even individuals in their everyday lives show inconsistencies that violate SEU, and so assuming that the state acted like a person would not save this theory.

Summary: Assuming people act as though they were maximizing utility will not answer many hard question.