Appendix C

AERIAL SEARCH AS A BASIS FOR VERIFICATION DECISIONMAKING

This appendix covers three major points:

1. Evidence gained from aerial search is best used to chose between competing hypotheses, not merely to bolster selected hypotheses. This is especially true of negative evidence.

2. The process of arms control treaty monitoring leads to a situation in which prior subjective judgments and a continuous influx of evidence, much of it negative, need to be blended into a single assessment. This need can be satisfied through the use of Bayesian, as opposed to classical, statistics.

3. Bayesian methods allow calculation of the likely benefit conferred by aerial monitoring, provided—among other things-that the harm inflicted by a violation can be expressed in the same terms, e.g, dollars, as the rest of the calculation. Though difficult, the assignment of a dollar cost to a violation is sensible if we are rationally to allocate money to forestall or detect such violations.

Discussions of unknown Soviet behavior often include statements of the form "We have no evidence of their doing that" on the one hand and "You have no way of knowing everything they're up to" on the other. *Many* times, each of these statements will be true (see box C-l), and the result a standoff. Yet it can be possible to get more mileage out of negative evidence than is often obtained. The trick is to use it to compare the relative likelihoods of competing hypotheses.

When using aerial surveillance (or indeed any means) to monitor an arms control treaty, the way to use the information it provides is to compare the answers to two questions: 1) "Given that the situation is as I think (or fear) it might be, how likely would I be to see what I am seeing?' and 2) ''If the situation were otherwise, how likely would I be to see what I am seeing *anyway?"*[1] Comparison of these likelihoods allows the evidence--even if it is merely the negative evidence provided by a flight that saw nothing-to mod@ one's assessment of the situation while avoiding the fallacy of "affirming the consequent.

Box C-1—Introducing an Example

The following example will be used in the next several boxes to show various approaches to the problem of sizing an aerial monitoring effort and interpreting its results.

Concern exists that 50 SS-20 transporter-erector-launchers (TELs) might still be somewhere in the Soviet Union.[1] In round figures, the Soviet Union could have about 5 million square miles of ground accessible to off-road-capable TELs: considering their ability to hide in woods, or to spend considerable portions of their time in shelters, the TELs might degrade the ability of the aerial monitoring aircraft to the point that it has a l-mile effective sweep width. Given the 5-million-square-mile area a l-mile sweep width of aerial monitoring aircraft for SS-20s, and a 1,000 mile flight path for each flight, as discussed inch. 6, each flight has approximately a

$$1 \times 50 \times 1000/5{,}000{,}000 = 0.01$$

or 1 percent chance of detecting some SS-20 or other.[2] Note that this figure is very much different from the flight's

$$1 \times 1 \times 1{,}000/5{,}000{,}000 = 0.0002$$

or .02 percent chance of detecting one SS-20 in particular.

[1] The 50-missile level has been cited as a ''ballpark'' threshhold of military significance for the SS-20. See *The INF Treaty*, hearings before the Committee on Foreign Relations, United States Senate, 100th Congress, 2nd Session, Part 3, p. 45.

[2] This approximation is valid only because the low density of targets allows us to disregard the effects of random ''clumpiness'' in their distribution. With a hundred times more targets we would not have a 100-percent chance per flight of finding at least one, because occasions on which we saw one would be balanced out by occasions on which we did not see any.

---

[1] Physicians are taught this adage for diagnosticians: "When you hear hoofbeats, don't think of zebras."

[2] In this fallacy, also known as post hoc, ergo propter hoc, the second question stated above is not considered. For example, the argument ''If the Russians were to launch a battleship, we would first see them build a large hull; we see a large hull, therefore they will launch a battleship" neglects the possibility that the large hull is being built for some other project, e.g., an aircraft carrier. Those who reason in this way can construe negative evidence almost any way they want to: ''If the Russians were building a nuclear airplane they would say nothing about it; they have made no mention of a nuclear airplane, therefore they are working on one. '

Box C-2—Probabilistic Approaches

The l-percent chance (see box C-1) of detecting some SS-20 or other remains unchanged from flight to flight; the treaty-violating TLIs are mobile and can be redeployed between flights, perhaps to regions previously inspected, so that there is no build-up of information about TLI-free regions. The sampling situation is thus one of *sampling with replacement.*[1] Each inspection has a l-percent chance of finding some launcher or other and a complementary 99-percent chance of not doing so. Taken together, each of a set of inspections must miss all of the targets if the whole set is to turn up no targets; their chance of doing so is 99 percent per inspection, so

1 -(1 - .Ol~*rOfm* = chance of finding a violation.

Many violation scenarios of concern are at base breakout scenarios, so one should consider the question of detecting a force during its installation, not afterwards as is normally done, If one assumes constant rates of inspection and violation-installation, one can equate the problem of finding one of 50 launchers as they are installed over a year to that of finding one of 25 launchers that are present the whole time.[2] Deployments and inspections might go on at the rate of one per week, so in the 50 weeks required to deploy the force the inspecting side will have a

$$1 - (1 - .O1/2)^{50} = 22 \text{ percent}$$

chance of finding the violation before it is complete.

Simple probability as well as intuition suggests that if each flight has a l-percent chance of finding a violation, something must be significant about 100 flights. The calculation that

$$1/0.01 = 100,$$

shows that, on the *average,* 100 flights will be needed before the first violation is seen.[3]

[1] The standard classroom example of sampling with replacement is a situation in which one has an urn containing unknown numbers of blue and red balls. The problem is to determine the proportion in which the two colors are present by repeatedly drawing out a single ball, examining it, throwing it backin, drawing another ball at random, and so on, In the present instance, considering red balls the violations, the question is how one's confidence that the urn contains no red balls increases with each blue ball examined.

[2] R. Scott Strait of Lawrence Livermore National Laboratory not only demonstrated this surprisingly simple fact but @O pointed out its importance inthe analysis of the ubiquitous breakout scenario.

[3] The probability distribution underlying this statement is the "geometric distribution."

Though often fruitful even when informally worded, the above approach can be codified mathematically using *Bayesian statistics. This* appendix will show some examples of how such calculations could be done; the search models in ch. 6 readily provide the needed inputs.

Experimental use of Bayesian statistics at the Central Intelligence Agency (including use in problems of photointerpretation) led some to conclude that the method held only limited promise.[3]

Asking the Right Question

Absent any detection of a violation, it is natural for decisionmakers to want to know how the negative evidence accrued up to any given present time affects the probability that an undetected violation exists. The natural tendency of analysts is to cast the answer in terms of the probability that a particular level of violation would have been detected if it existed (see box C- 2): high levels of this probability are taken (given that no violation has been found) as providing high confidence that the particular level has not been exceeded. (See box C-3.) Yet decisionmakers want to know something else, namely the probability that a specified level of violation actually exists, given the evidence collected.

The split can be seen as dividing the analysts and the decisionmakers almost along the classical-v.-Bayesian split in statistics itself. The analysts' view of the problem corresponds closely to that of the classical statistician (who believes in a single-though unknowable--reality and views experimental data as random samples thereof), while the decisionmakers' view more strongly resembles that of a person working within the Bayesian paradigm (who is willing to assign probabilities to various "realities" and views experimental data as fixed "givens"). (See boxes C-4 and C-4a)[4]

A flaw present (though not unavoidable even under the classical paradigm) in many attempts to quantify matters of verification arises from shifting assumptions about the size of a violation. The classical analysis prevalent in the verification literature scales the monitoring effort on the basis that a violation will consist of a militarily significant

[3] See Walter Laqueur, *A World of Secrets* (New York, NY: Basic Books, 1985): pp. 299-302. Laqueur cites Richards J. Heuer, *Quantitative Approaches to Political Intelligence: The CIA Experience* (Boulder, CO: Westview Press, 1978).

[4] Repeated instances of this sort of split and a renewed emphasis on what might be called a decision-analytic view of statistics have led to a sharp increase in the teaching of Bayesian statistics, once greatly out of favor in academia.

Box C-3—The Classicist's Approach

The classiccal statistician is used to dealing in high (90 percent, if not 95 percent or 99 percent) confidence levels, and to looking at data and making statements about the state of the world. The statistician will want to assess an inspection scheme by standing in the shoes of somebody examining some large amount of inspection data after the flights have been made; the decisionmaking question is one of interpreting-the continuing flow of reports that no treaty-violating-missile launcher has been sighted.

Mom specifically, the classicist seeks to make a statement about the world to which he or she can assign a high probability that it is correct. In the present case, for example, a classicist might like to be able to make the statement that "there is no significant violation" and assign to that statement a 90-percent chance of being correct. The classicist would therefore request that 229 flights be made and when these flights reported that no violations had been seen, the classicist would make a finding of compliance "at the 90-percent confidence level."[1]

The basis for this statement is the binomial distribution, according to which there is a 10-percent chance that 229 flights would all miss the violation even if it were present in its smallest militarily significant form, given a l-percent chance per flight of finding it, i.e. that

$$0.9 = 1 - (1 - 0.01)^{229}$$

so that 229 flights would have a 90-percent chance of detecting the militarily significant violation if it were present.[2] Two hunched and twenty-nine flights which find no violation are required to make the classicist "reject the null hypothesis" that a significant violation exists.

The classicist's statement that he or she has 90-percent confidence in the proposition "no significant violation is present" (given 229 flights which found no violation) might lead a decisionmaker to think that there is a 10-percent chance that a significant violation exists. We may hope that the decisionmaker would ask the classicist to elaborate upon this point. The answer would be that the classicist feels "confidence in the statement at the 90-percent level" because the statement has at *least* a 90-percent chance of being true: it was produced by a process,[3] which, if applied repeatedly, would yield correct statements at least 90 percent of the time. When no violation is present the statement will always be made and will always be right: when a violation is present the opposite statement (a finding of noncompliance, certain to be correct because it is backed by evidence from at least one flight) will correctly be made 90 percent of the time[4] and an erroneous finding of compliance will be made 10 percent of the time. Thus the classicist feels that he or she "bats 1,000" in nonviolation situations and "bats 900" in violation situations, and is therefore justified in claiming to bat at least 900 (i.e., make "statements at the 90-percent confidence level") *even without knowing the true mix of violation and non-violation situations. The* classicist might also point out that his or her method will never result in an erroneous finding of non-compliance-findings of noncompliance are only made with hard evidence in hand.

Ninety-five Percent confidence could be obtained by flying 298 flights instead of 229. The correct choice of confidence level depends upon the perceived balance between the cost of the flights and the cost of lack of confidence.

The classicist decries the Bayesian approach as "subjective," because, among other reasons, the Bayesian assigns a *probability* to the existence of a militarily threatening treaty-violating missile force. The classicist views such an assignment as incoherent because the force either exists or it doesn't.

---

[1] If a violation is found, the classicist will be able to make a finding of noncompliance at the 100-percent confidence level.

[2] More generally,

$$\text{confidence} = 1 - (1 - \text{detection probability})^{\text{inspections}}$$

This easily-derived equation pervades the verification literature, being used, for example, by Dunbar Lockwood in "Verifying MART: From Satellites to Suspect Sites," *Arms Control To&y,* October 1990, p. 16.

[3] "The classicist is absolutely opposed t. the interpretation that the 90 percent refers to the probability that the true popualtion mean lies within the specified interval .... A classicist is willing to generate statements with respect to the probability that *the procedure* leads to the generation of correct statements ...." (Robert Parsons, *Statistical Analysis: A Decision-Making Approach (New York, NY:* Harper and Row, 1978): p, 329, emphasis in original.)

[4] Note that this figure corresponds to the threshold level of violation. Larger violations would have had a single-flight detection probability of more than 1 percent and thus would have an overall chance of more than 90 percent of being noticed by the 229 flights.

number of treaty-limited items (TLIs), e.g., 50 missile transporter erector launchers (TELs), and is prepared to return a verdict of "not guilty" (at, say, the 95-percent confidence level) if that effort reveals no TELs. Yet the classicist—like anybody else—is prepared to return a "guilty' verdict if even a single illicit TEL is found. The classical paradigm precludes the direct translation of this finding into an assessment of the probability that a force of 50 missiles is present. In contrast, the Bayesian's method permits such a finding; if a missile or two are seen, the Bayesian will of course announce a violation but will also continue his or her estimation efforts, updating the probabilities that 0,5, or 50 illicit launchers have been fielded. (See box C-5.)

Box C-4-The Bayesian's Approach

In contrast to the classicist (see box C-3), the Bayesian analyst declines to make flat statements, even at specified levels of confidence, preferring instead to describe the number of missiles present in terms of a probability distribution. He or she views the classicist as "concealing information"[1] by boiling everything down to a statement regarding the presence or absence of 50 missiles and a figure for a reliability of the process that generated the statement[2].

The Bayesian addresses the aerial monitoring problem by assigning probabilities to the existence of various numbers of treaty-violating missiles and goes so far as to insist that even before the inspections there must be some a priori or "prior" distribution for this number. He or she then interprets the flight data in light of how likely each fight's outcome would be given each number of missiles. Finally, the Bayesian combines the prior probabilities with the probabilities that each flight would turn out as it did, arriving at a "posterior" set of probabilities of the existence of the different numbers of treaty-violating missiles.

To implement this approach, the Bayesian needs a single-flight probability of violation-discovery for each number of missiles. Continuing with our example, we will recall the 0.02-percent chance of detection per missile. The Bayesian also needs a prior distribution of different levels of violation. Suppose, for sake of illustration, that somebody supplies a prior distribution to the effect that there is a 5-percent chance that the Russians have a threatening 50 extra missiles, a 70-percent chance that they have a token 5 extra missiles, and a 25-percent chance that they are not cheating at all, This prior, perhaps reflecting a belief that the Soviets are trying to comply but have not quite tracked down all their missiles yet, is expressed by the leftmost two columns of the table, "Missiles" and "Prior."

Missiles	*Prior*	*Seen*	*Joint*	*Renormalized*
0	*25%*	100%	*25%*	*31%*
5	*70%0*	*80%*	*56%*	*69%*
50	5%	10%	1%	1%
	100%			100%

Now suppose that after 229 flights (not that 229 is a particularly special number of flights for the Bayesian) no violations have been seen. The Bayesian generates the "Seen" column by computing the likelihood of this result under each of the three assumptions about how many missiles are actually present. The probability that no missiles would be seen under an assumption that no violations exist is, of course, 100 percent. If 5 violations are present, there is a O.1-percent chance per flight of seeing at least one and an 80-percent chance of seeing at least one in 229 flights. As arose in the classical case, 229 flights have a 10-percent probability of leaving 50 missiles totally unnoticed. The fourth column shows the joint probabilities of 0,5, or 50 missiles being in existence *and* resulting in zero sightings after *229* flights. The entries in the fourth column are the products of the respective entries in the second and third columns because of the chain rule of probability: the probability of events A and B occurring jointly is the probability of A multiplied by the probability of B given that A has occurred. The fourth column thus shows, for each number of missiles, the probability that they are present multiplied by the probability that the flights would have the result that they did (no sightings) with that number of missiles present. These probabilities do not add up to 100 percent because there was not a 100-percent chance that the flights would see no missiles. The fifth column shows the entries of the fourth column renormalized, (that is, the entries are divided by their sum) so as to make their total 100 percent as is required of a probability distribution.[3]

Thus the Bayesian's report on the results of 229 flights shows that the probability of 50 existing undetected launchers has been greatly eroded by the negative evidence of the 229 flights, while the probability that 5 launchers exist undetected has only gone down slightly, because 229 flights are not enough to prove very much about such a small force. The clean bill of health from the 229 flights has somewhat bolstered the case that the other side is not violating the treaty at all.

1Barry Blechman, "The Impact of Israel's Reprisals on Behavior of the Bordering Arab Nations Directed at Israel," *Journal of Conflict Resolution,* vol. *XVI, No. 2,* June 1972, p. 169.

2He or she sees the classicist's objection to the probabilisitic description of a *fait accompli* (albeit an unknown one) as a philosophical quibble: people buy and sell unscratched lottery tickets on the view that the tickets have some probability of winning, when in fact each ticket is foreordained to either win or lose.

3 This step is needed because we have been considering only certain possible numbers of violating missiles (0, 5, and 50), not all possible numbers (O, 1,2,3,.. .). The fact that the original probabilities (of 0,5, and 50 missiles being present) summed to 100 percent was somewhat artificial in the first place: the originator of such an estimate would ensure for appearances' sake that the sum was 100 percent but would not mean thereby to exclude the possibility of 49 missiles being present. The mathematical operations do not carry this artificialitythrough, so it has to be reintroduced via renormalization. Additionally, the fourthcolumn, being a joint probability, cannot sum to more that the marginal probability of either of its component events (columns 2 and 3). This will always be less than 100 percent if the evidence collected provides any information at all. Some would object to the blending of the renormalization step (required only because the prior did not consider all possible numbers of missiles) with the division of the joint probability (column 4) by the marginal probability of the evidence. As a practical matter, however, the two steps are done at once by the same division. The marginal probability of the evidence is the denominator, called the "prepmterior," in Bayes' Rule. The fifth column is thus the posterior probability of the violation in the first column given the observed evidence—no violations sighted.

The Bayesian's report has the virtue that it answers the question the decisionmaker wants answered: "What is the probability that the other side is violating the treaty?" The Bayesian requires, however, a prior distribution in whose light he or she can consider the findings of the aerial monitoring flights. While the classicist might lodge an objection that the Bayesian is relying on subjectively obtained prim information, the Bayesian can retort that the classicist's 50-missile theshhold and 90-percent confidence level were in themselves subjectively obtained prior standards-the Bayesian approach uses no such standards.

Moreover, the Bayesian could continue, the classicist's standards are inconsistent. The requirement for 229 flights stemmed from a need to determine (with 90-percent confidence) that the Soviets had not fielded a force of 50 illicit missiles. A clean bill of health after these flights will result in a finding (at the 90-percent confidence level) that the Soviets are not in violation; the classicist has retained faith in the 229-flight figure while rejecting the 50-TEL figure on which it was based.[4]

[4] Arnold Zellner makes a more general version of this point (paraphrasing E.T. Jaynes): "if the null hypothesis is rejected in a non-Bayesian analysis, then so too is the distribution of the test statistic tbatled to the decision rule for rejection." ("'Bayesian Inference," *Time Series and Statistics (New* York, *NY:* W.W. Norton, 1990), p. 56.)

Box C-4a-The Effect of a Different Prior

As a practical matter, the Bayesian may have difficulty mustering a prior distribution which is acceptable to all concerned.[1] With a somewhat different prior, a somewhat different set of posterior probabilities will emerge from the same flight data, Let us examine the effect of starting with a different prior distribution, perhaps created by a different analyst:

Missiles	*Prior*	*Seen*	*Joint*	*Renormalized*
0	20%0	100%	20%	32%
5	50%0	80%	40%	63%
50	30%	10%	3%	5%
	100%			100%

Again, the flights have seen no violations, but this analyst's calculations give far more weight to the possibility that a significant violation exists because of his or her use of a different prior. The Bayesian will point out that as more information comes in, the importance of the prior is diluted. Indeed, the right-hand column shows that the evidence of 229 flights that sight no transporter-erector-launchers results in almost the same set of posterior probabilities with the second prior as it did with the fret.

The alternative prior distribution shown above is in some sense a greater expression of ignorance than the original prior, in that it accords more nearly equal probability to the three cases. One might be tempted to evade the responsibility of creating a prior by simply assigning equal probability to all the possibilities.

The difficulty with this scheme[2] lies in listing "all the possibilities. " One could start with zero missiles and count up, but surely there is some upper limit to the number that the Soviets could deploy, and surely some values near that limit are less likely than some lower ones.

Worse, even if one could list all the possibilities, it is far from clear how one should assign probabilities to them in order to reflect an absence of preconceived notions. For example, assume that incontrovertible evidence (perhaps related to production capacity) shows that there can be no more than 99 illegitimate missiles deployed. An analyst could then perform calculations like those shown above, according a prior probability of 1 percent to each of the 100 possible numbers of missiles (O, 1,2,3, .. .99) and defending this prior on the grounds that its uniformity reflects the absence of any preconceived notion as to the number of missiles. However, the assigned l-percent probability that zero missiles have been deployed reflects a 99-percent certainty that the Soviets are cheating in some degree, hardly an absence of preconceived notions.

In any case, to seek parsimony of assumption through the use of an ignorant prior is to underestimate the amount of information available size of a likely treaty-violating force is bounded from above by economic considerations and from below by military effectiveness.

[1] A difficulty mentioned by D.V. Lindley in "Statistical Inference,'' *Time Series and Statistics,* op. cit., box C-4, p. 291).

[2] Termed ''notorious'' b, Koopman. (Bernard Osgood Koopman *Search and Screening,* (New York, NY: Pergamon, 1980), p. 286.)

Aerial Monitoring as a Basis for Confidence

We should think of aerial monitoring in a broader sense than that of simply catching Soviet violations of arms control treaties. (Box C-6 sets the stage for the coming examples.) Aerial monitoring might extend beyond treaty monitoring to the more general function of providing assurance that the other side was not mobilizing for war. Even with regard to treaty monitoring the function of providing confidence that the treaty is not being violated would be at least as important as the function of giving warning that it had been.

To perform this compliance-monitoring function, however, the aerial monitoring regime must bolster confidence in a negative-the proposition that the Soviet Union is not violating a treaty or mobilizing for war. As experienced lawyers and debaters are well aware, to "prove a negative" can be difficult or impossible.

Treatment of Negative Evidence

As part of the Intermediate Range Nuclear Forces (INF) Treaty verification process, the Soviets gave the United States a data package regarding the SS-20 intermediate range ballistic missile. Asked if the package was accurate, Admiral William J. Crowe responded, "We do not have the evidence or the conviction to say that it is inaccurate."[5] This "Scotch verdict" is informative only if one has some idea of how likely the United States would be to possess information disconfirming the package if the package were in fact inaccurate.

Asked about many possible scenarios, intelligence analysts will respond, "We have no evidence of such activity." These analysts can hardly be faulted for declining to speculate in the absence of evidence, but one must exercise care in interpreting their silence because it raises the question, "Would you have evidence if they were actually doing it?" This question gets very close to the Bayesian's question, "How likely would I be to see what I am seeing if the activity were going on?"

Much of the utility of the Bayesian formulation lies in the fact that it can incorporate negative evidence as well as positive evidence. A flight that produces no evidence of treaty violation is viewed by the Bayesian as supporting the case for treaty compliance *insofar as* a violation, were one to exist, might have been noticed by the flight.[6] This measured use of negative evidence differs from the naive conclusion that "absence of evidence is evidence of absence" and from the traditional conclusion that a lack of evidence proves nothing either way.

Discussion of treaty monitoring begs the question of what action would be undertaken in response to the discovery of a violation, so such an assessment must hinge upon the action that the United States would take upon finding an anomaly or violation as the result of a flight, and upon the difference in impact upon the United States of finding an anomaly or violation sooner rather than later.

This report will not attempt to prescribe any such actions, but we will consider one-the deployment of countervailing intercontinental ballistic missiles (ICBMs) —as an example of how the costs of reacting to a violation affect the monitoring process itself.

Harm Inflicted by Treaty Violation

It is difficult to assess, let alone quantify, the harm inflicted by a treaty violation. (Boxes C-7 and C-8 show two possible approaches.) Much depends upon the treaty, and the violation, but even so one might credibly follow any of several lines of reasoning and attach costs varying from close to zero to close to infinity to the harm inflicted by a given treaty violation. Why then even try to estimate the cost imposed by an arms control treaty violation? In particular, the idea of assigning a dollar cost to the harm caused by deployment of treaty-violating TLIs may seem bizarre. However, we budget treaty-monitoring resources in dollars. Therefore, we need some estimate of what these dollars buy us in terms of harm avoided, just as if we were instead budgeting for the construction of a retaliatory force or a system of active defenses.[7]

The Expected Value of Information

For the purposes of illustration, we will proceed with an analysis based on the example in box C-8, of $500 million per undiscovered (and hence encountered) missile and $200 million per discovered (and countered) missile. Different assumptions will produce different results.

Let us compare the cost of performing the aerial monitoring inspections v. the cost of not doing so. The reader is again reminded that this calculation is an

---

[5] *NATO Defense and the INF Treaty*, Hearings Before the Committee on Armed Services, US, Senate 100th Cong., 2d Sess., part 1, p. 121. Asked to clarify the point regarding "conviction," the Admiral replied "I do not believe that the evidence supports it in what we have."

[6] More pessimistically, a flight that produces no evidence of a treaty violation can be viewed as evidence that the collection equipment isn't working, insofar as one thinks that violations are present, The methodology presented in this chapter coulbe expanded to characterize formally this updating of a prior probability that the equipment doesn't work. A closely relatedifficulty is that one might be using the equipment to look for the wrong thing, as may have been the case in the anti-Scud campaign of the Gulf War, in which the search for Scud TELs overlooked the more prevalent "mobile erector launchers' expediently produced in Iraq.

[7] Classical statistics would require the even more problematical imputation of value to *confidence* in missiles *not being present*. While One might reasonably think that the presence of 50 illegitimate missiles isapproximately twice as deleterious as the presence of 25 (or the 50-percent probability of the presence of 100), there is little intuitive appeal to the idea that the absence of 100 missiles is twice as nice as the absence of 50.

Box C-5—The Bayesian's Report After a Sighting

The table shows what the Bayesian would report if one transporter-erector-launcher (TEL) were sighted in the course of 229 flights, given the same prior distribution for the number of illicit launchers deployed as was used in box C-4.

Missiles	*Prior*	*Seen*	*Joint*	*Renormalized*
0	25%	0%	0.00	0%
5	70%	18%	0.13	92%
50	5%	23%	0.01	8%
	100%			100%

The Bayesian again generates the "Seen" column by computing the likelihood of this result under each of the three assumptions about how many missiles are actually present. The probability that one missile would be seen under an assumption that no violations exist is, of course, O percent: sighting a TEL would be impossible if there weren't *any*,[1] so no Chance remains that the Soviets have fielded no illicit TELs. If 5 were fielded, 229 flights would have an 18-percent chance of seeing exactly one; if 50 were fielded, 229 flights would have a 23-percent chance of seeing exactly one.[2] Combined with the respective prior probabilities for the deployment of these numbers of missiles, these probabilities result in a 92-percent chance that 5 TELs are deployed and only an 8-percent chance that a significant violation featuring 50 illicit TELs is underway.

[1] Though *thinking* that one had sighted a TEL would not be. A more complete analysis would take into account the possibility of "f&R alarms' in which TELs are reported where none exist.

[2] It may seem paradoxical that a tenfold increase in the deployment causes such a modest increase in the probability ofseeing exactly One TEL. The reason is that although 1 TEL is a lot to find if 5 are deployed, it isa small number tofind if 50 are deployed-the sighting of exactly 1 TEL is in fact mild evidence *against the* proposition that 50 are deployed, because one would expect the searches to reveal more than 1 TEL in that case.

Box C-6—Introducing Another Example

The following example describes our treatment of cost considerations. We will maintain our focus on a violation that would clearly be of some harm to the United States, namely a violation of a ballistic missile launcher limit. The following assumptions, chosen purely for simplicity and plausibility, will govern our consideration of the problem: Aerial monitoring flights detect illegitimate missile launchers as described in box C-5: each flight exerts an effective sweep width of one mile over a 1,000 mile path in a 5-million-square-mile-region of the Soviet Union. The actual search portion of a flight consumes 2 hours of flying time, with another hour spent in takeoff, landing, gaining altitude, and other nonsearch portions of the flight. The flights cost $50,000 apiece all told. This estimate is based on a 6-hour Washington-Los Angeles flight breaking even at about 200 passengers paying about $300 apiece. Flying the aerial monitoring aircraft will cost roughly $10,000 per hour for 3 hem. The other $20,000 is the cost of the reconnaissance functions themselves, including interpretation performed afterwards. To check this estimate, we may note other indicators of the cost per flying hour of various types of aircraft. The manufacturer of a twin-engine turboprop aircraft cites its $2,000/hour overall expense as a compelling factor in its favor.[1] Aerobureau, a private firm, proposes to lease a Lockheed Electra (the civilian equivalent of a P-3 Orion), equipped with "side-looking airborne radar, infrared and low light television sensors,' to TV news networks for a $250,000 6-month lease and a $2,000/day operating fee covering the aircraft and flight crew.[2] The owner/operator of a USAAF B-29 restored to flying condition reported a $3,000/hour cost of operation with a volunteer crew.[3] A B-1 wing flying four sorties per plane per month expends $242 million in operating costs per year, suggesting a cost per sortie of almost $100,000.[4]

Costs of different aerial photography systems are often usefully compared in terms of dollars per square mile.[5] Assuming that the l-mile sweep width cited above stems from a 20 percent effective photographic search of a 5-mile swath (not an unreasonable assumption: see box 6-A), this search costs $10 per square mile.

Some number of treaty-violating launchers may be deployed. The U.S. side's set of prior probabilities for the number of illegitimate missiles deployed is the same as it was in box C-5.

[1] John King, "Airborne Remote Sensing for Open Skies" *Open Skies* (Toronto: York University Center for International and Strategic Studies, 1990), p. 33.

[2] Aviation Week and Space Technology, August 20,1990, p. 13.

[3] Personal communication.

[4] Testimony of General John T. Chain Jr., . *Threat Assessment; Military Strategy; and Operational Requirements,* hearings before the Committee on Armed Services, U.S. Senate, Mar. 7, 1990: page 898.

[5] See, for example, Amrom H. Katz, "Let Aircraft Make Earth-resource Surveys,' *Astronautics and Aeronautics,* June 1%9, reprinted under the same title as RAND Paper P-3753, available from The RAND Corp., Santa Monica, CA. Katz's costs per square mile are considerably lower than than the figure presented here not only because of the greater value of the dollar in 1969 but also because of the less-demanding resolution requirements of the earth-resource survey mission.

Box C-7—Harm Inflicted by Treaty Violation: Valuation According to Extra Potential Damage

One could also place a dollar cost on the harm done by a violation accordingto the destruction the violating weapons would inflict were they used against the United States. If the extra loss of life and limb--always difficult to cast in dollar terms--were costed at the rates paid in injury liability cases, the casualties from one nuclear missile could soar far into the billions. On the other hand, densely populated areas are presumably already targeted by the missiles and other strategic nuclear delivery vehicles allowed under arms control treaties: treaty-violating missiles will be applied to marginal targets left uncovered by the legitimate force.

The damage done to inanimate objects is somewhat less difficult to cast in dollar tarns. One estimate of this amount comes to 36 times the cost (procurement and lifetime operation, or about $200 million for the Peacekeeper) of the weapons themselves,[2] so that a violating missile imposes a cost of about $7 billion on the United States. The factor of 36 could be assailed on the grounds tha t although it assumes complete obliteration of the United States, its imputation of worth is restricted to "asset value," and as such includes neither loss of life and limb, or loss of items and sites of cultural value. Thus the factor of 36, obtained by assuming that the entire heavy Soviet intercontinental ballistic missile (ICBM) force would destroy the entire United States, may be assailed as too high on the grounds that it is an average value, not a marginal value.

[1]If not directly, then through the legitimate missiles they displace from the higher-priority targets.

[2]Gregory Canavan and Edward Teller (in "Strategic Defense in the 1990," Nature, Apr. 19, 1990) derive this estimate for the SS-18, noting that "overestimates of the effects of nuclear explosions may have caused many readers to guess a higher figure." They take the SS-18 to cost as much as its U.S. counterpart, the Peacekeeper, and they assume that the entire force of 308 SS-18s could destroy the entire United States. The value thus destroyed is derived from U.S. gross national product and economists' usual ratio (about 4:1) of total assets to production. An upper limit on the total worth of American people, assets, and cultural items that could possibly be destroyed in a war would be the total insured value underwritten by all insurers, property, casualty, and life. People, vehicles, and land aside, insurers cover about $12 trillion in assets: inclusion of vehicles and allowance for the existence of various forms of uninsured property would raise this figure to $15 trillion or more.

Though we maybe startled by the seeming lowness of the 36:1 ratio of the missile's cost to its destructiveness, we should recall that conventional bombing notoriously destroys less than it costs.

illustrative one based on the assumptions listed above. Moreover, the real issue is the effect of aerial monitoring upon monitoring as a whole; the correct comparison would really be all other monitoring with versus without aerial monitoring added. Without any aerial monitoring flights, the cost is the harm done by the presence of any illegitimate missiles, expressed in dollar terms as explained above. Using a prior distribution for the probabilities that various numbers (including zero) of illegitimate launchers are present, we may find the expected value (translated into millions of dollars) of the harm done by the missiles: $3 billion.

Missiles	*Prior*	*cost*	*E(cost)*
0	25%		
5	70%	$ 2,50:	$1,750:
50	5%	$25,000	$1,250
			$3,000

Aerial monitoring flights would have some chance of detecting the violations committed in the 5-missile and 50-missile cases. Two hundred flights, each effectively sweeping a l-mile swath 1,000 miles long, might each have a 0.02 *percent* chance (as we have seen in ch. 6) of finding one launcher: this probability can be multiplied by 5 of 50 to get the single-flight probability of finding a violation, and the chance of not finding the violation then raised to the 200th power to obtain the probability of *not* finding a violation during the 200 flights. We may thus compute the probability of finding the violations. Recalling that we assume that a violation costs $200 million per launcher to counter and $500 million per launcher if it remains hidden, recalling that the 200 flights themselves cost a total of $10 million, and using our prior probability distribution for the chances that each level of violation (including none at all) is present, we may find the expected value of the cost of each violation

Missiles	*Prior*	*Find violation*	*E(cost)*
0	25%	0%	
5	70%	18%	$1,564
50	5%	87%	$ 601
			$2,170

For example, the 5-launcher violation has an estimated 70 percent probability of being present in the first place and a probability of

$$1 - (1 - 5 \text{ X } 0.0002)^{200} = 0.18$$

of being detected by some one of the 200 flights or other if it is present, so that the total expected cost of the 5-missile case (including the cost of performing the inspections, even if they find nothing), is

$$0.7 \text{ X } ((5 \text{ X } (0.18 \text{ X } 200 + 0.82 \text{X } 500)) + 10) = 1564^{8}$$

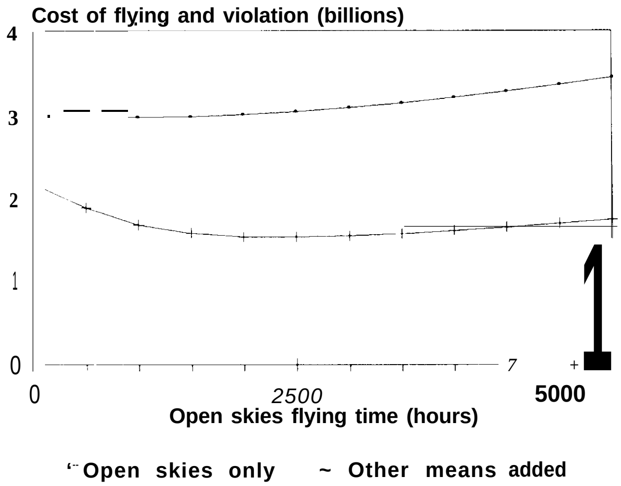
or $1.564 billion: $200 million per missile if the violation is detected and $500 million otherwise. The expected cost of the situation as a whole is $2.18 billion if inspections are undertaken-less than the cost without the inspections. Therefore the inspections are the preferable course of action under these assumptions, though by a small enough margin that other reasonable assumptions could

Box C-8—Harm Inflicted by Treaty Violdation: Procurement Cost of Counteracting a Violation

Another position would evaluate the violation at the cost of the violating weapons, or of their U.S. equivalents, on the grounds that the U.S. response to a violation would be to build similar weapons to maintain parity. However, the U.S. preference for arms control over arms racing suggests that the United States feels that having weapons aimed at it has a cost greater than the cost of aiming our own weapons back. Very roughly, then, this middle view might attribute to an encountered treaty-violating missile a one-time cost of $500 million to the United States.

A violation which was found and countered by the deployment of a U.S. missile can be more readily evaluated at the cost to counter, roughly $200 billion for a U.S. intercontinental ballistic missile (ICBM). Use of other legs of the triad to counter a violation would cost. more: ICBMs offer the least cost per alert warhead.

Figure C-l—Notional Costs of Inspections and Violations (detected and undetected violations)



The expected sum of the cost of the inspections and the imputed costs of violations represents the total cost to the Nation. An amount of flying can be chosen to minimize this total. The upper curve depicts a situation, described in the text, in which only Open Skies flights can detect a violation. The lower curve shows the contribution of some other means of detection that, albeit at a cost, greatly increases the probability that a violation will be detected.

SOURCE: Office of Technology Assessment, 1991.

lead to the opposite conclusion.[8] One unstated assumption with which serious issue might be taken is that finding one or more illegitimate launchers somehow brings about discovery of the whole violation. This is one effect which brings down the cost of the 50-missile case. (Another is the low prior probability assigned to that case.) The assumption is not totally unreasonable, in that discovery of a violation would shift the whole treaty-monitoring effort into high gear and probably bring other intelligence collection assets to bear. The expense of using those extra assets should properly be charged against the cost of countering a treaty-violating missile, perhaps increasing the $200-million figure we have been using. Another unstated assumption has been that the aerial monitoring flights afford the only opportunity for catching the violations. In the real world, this assumption is unwarranted and the analysis should be redone using correct overall costs of monitoring and correct overall probabilities of catching violations.

Savings Expected From Aerial Monitoring

Referring to the computations of the previous sections, we may attempt to assess the worth of the 200 aerial monitoring flights by finding how much money we can expect them to save us. Given all of our assumptions, including the prior distribution of the probabilities of the various violations, we find an expected cost of $3 billion without the flights and one of $2.17 billion with them so that the flights save us $830 million net, including the costs of the flights.

More generally, we may find the worth of various numbers of aerial monitoring flights by finding their expected net cost and comparing it to the expected $3 billion cost of not having any flights at all, always keeping in mind that the result depends upon all of our assumptions, including our *a priori* estimates of the likelihood of various levels of violation and our dollar-cost characterizations of the harm done by countered and encountered violations.

8 Certainly the precision of some of our assumptions does not justify the retention of four decimal places in the result, as has been done for the sake of clarity in illustrating the structure of the calculations.