INSPECTING FOR WEAPONS PRODUCTION

A MODEST COMPUTER SIMULATION *

By

J. DAVID SINGER and HIROHIDE HINOMOTO

Mental Health Research Institute,
University of Michigan

Department of Industrial Administration,
University of Illinois

1. Introduction

During the post-war disarmament nego-
tiations between the two major powers,
one of the most intractable issues has been
that of international inspection. In general,
the Soviets have preferred a minimum of
inspection, instituted as late in each stage
as possible, while the West has pressed
for both a greater degree of inspection
and its institution as early as possible in
each stage of the complex arms reduction
process. Despite some modest progress on
principles, no concrete agreement on
specifics has yet been negotiated; even the
limited test ban treaty relies solely upon
national inspection procedures, carried on
outside one another's borders. And while
a break in the inspection impasse would
by no means suffice for the establishment
of a meaningful arms reduction arrange-
ment, there is little doubt that it never-
theless remains a necessary precondition.[1]

As the negotiations have proceeded in
their atmosphere of ambivalence and
acrimony, the responsible policy-makers
and representatives have devoted consider-
able effort to persuading their opposite
numbers and their domestic publics of
the reasonableness of their respective po-
sitions. But whether the audience is
addressed across the bargaining table or
via the mass media at home, the would-be
persuaders have labored under several
serious handicaps. Of these, perhaps one
of the more crucial has been a lack of
specificity based either upon experience
or upon experiment. Even in the matter of
underground nuclear test detection, the
credibility of the experiential or experi-

mental base has been seriously reduced by
the misleading interpretations placed upon
the resultant data. It seems safe to predict
that, until a fairly reliable — and credible
— body of data becomes available, the
adversaries will debate in vain; and even
if the adversary is persuaded, the home-
front hard-liners will still find it easy to
charge a 'sellout'.

If ex cathedra argument is to give way to
a measure of empiricism, and if very little
reliable experience is available, some sort
of experimentation would seem to be in
order. Such experimentation may take a
variety of forms, each with its peculiar
strengths and weaknesses. At the micro-
system level, for example, the investigator
might try to ascertain the effectiveness of
varied training programs by measuring
the accuracy and rapidity with which
inspection personnel can distinguish cer-
tain salient configurations or procedures;
or of the efficacy of various types of instru-
mentation for ascertaining the absence,
presence, or quality of specific chemical
or metallic substances. And at the macro-
level, we might experiment with varying
inspection grid patterns in order to
appraise the thoroughness, speed, and
economy with which a surveillance team
might cover a large geographical region.
In either case, models based on theoretical
considerations are put to the test in situ-
ations as close as possible to the real
world in which the inspection system may
be expected to operate. In addition, these
experiments may occur either in the field
or in a laboratory-type setting. The field
or factory setting has the advantage of

greater realism, but in the laboratory one is able to manipulate each of the key variables with a greater degree of flexibility.

As both the quantity and quality of research on inspection for disarmament increase, in government as well as out, the desirability and possibility of going into greater detail likewise increase. Just a few short years ago, there was almost no available unclassified literature on the subject; the journals prior to 1960 had carried a handful of papers and most of these dealt with rather general problems. Outside of the work dealing with detection and identification of nuclear weapons tests, the major detailed studies then available were those of Feld, Henkin, and Melman.[2] But the upsurge of serious work on the inspection problem has now made it mandatory that we become increasingly specific in our research.

In the absence of actual experience in using various inspection methods currently available, the rational selection of a particular method can only be based upon their anticipated relative effectiveness in a given range of situations. Such effectiveness may be appraised by: 1) abstract analysis, 2) laboratory or field experiment, 3) digital computer simulation using abstract (mathematical and logical) models, or 4) any combination of the above.

Unfortunately, too little of such experimentation has occurred, particularly in the matter of inspecting for clandestine production of proscribed quantities or types of weapon or hardware.[3] Given this situation, we have attempted a modest experiment, or computer simulation, based on a rather simplified set of production, evasion, and inspection models. Though the simulation might as readily be applied to any weapon type, we have thought in terms of one of four possible agreements regulating aircraft and rocket production. In matrix form, the four possibilities might be as follows, with a tentative ranking in order of inspection difficulty:

	Specified sub-classes	All weapons of a class
Limited Production	1	2
No Production	3	4
		(Least difficult)

The sort of ban examined in this study is type 3, which permits unlimited production of certain classes of airplanes and rockets and prohibits *any* production of other classes.[4] Aircraft which are capable of carrying and launching mass destruction weapons, including stand-off missiles, are prohibited, and the distinguishing characteristics of illegal planes are assumed to be: (a) the presence of bomb-bays, bomb shackles, and launching-guiding equipment; and (b) the absence of pressurization and seating provisions for passenger craft, and loading-unloading and storage provisions for cargo craft. As to rockets, we assume that those designed for delivering nuclear or conventional warheads over any range are prohibited, and that illegal rockets are distinguishable by: (a) the presence of ground-to-ground guidance mechanisms, re-entry cones, and payload fittings designed for warheads rather than non-lethal instrumentation; and (b) the absence of destruct mechanisms and provision for separation of instrumented payloads. In both the airplane and rocket category, we assume that a number of other characteristics exist by which legal production may be differentiated from illegal production; the particular criteria are, in any event, not crucial to the design of this type of experiment.

Given these assumptions, what is the purpose of an inspection system? Objectively speaking, it is that of reducing (but not completely eliminating) the probability of an evasion going undetected. As the probability of successful evasion goes down, two important consequences accrue. First, the potential evader (let us label him A) will — assuming the costs to him of being detected are high enough — be less motivated to evade. That is,

a rational and amoral group of national (or local) decision-makers will have a temptation to evade which diminishes as (a) probability of detection increases, (b) costs to them of being detected increase, (c) gains to them of successful evasion decrease. Second, as the potential evader's (A's) probability of successful evasion decreases, so will the other party's (B . . .N) *belief* that evasion is being attempted in the first place, all other things being equal. This tendency, in turn, will reduce B's temptation to evade, inasmuch as one of the factors which increases the temptation to evade is the fear that the other may be gaining a strategic advantage by engaging in clandestine evasion.

However, the efficacy of an inspection system cannot possibly be measured solely by the probability with which it will (or frequency with which it does) detect attempted evasions. Equally crucial is the degree to which it minimizes the probability of *false* 'detections'. The agreement could as readily be torn asunder by a high number of false alarms as by a high number of successful evasions. Thus, the effectiveness of the system must be seen as varying directly with the probability of detection *when evasion is occurring* (correct detection) and inversely with the probability of an alarm (false detection) when evasion is *not* occurring.[5]

Finally, no inspection system can be appraised purely in terms of its performance. Associated costs must be considered, and among these, two stand out as crucial. One is the simple item of monetary and material costs of installing and operating the system, and this is susceptible to relatively simple calculation. But the other is political onerousness, and lends itself much less readily to such appraisal and prediction.

2. *The inspection situation*

Given these general assumptions regarding the usefulness of the type of experiment proposed and the basic pur-

poses of a production inspection system, let us now describe the hypothetical inspection situation in moderate detail. Following that, we will move on to a presentation of the simulation procedure.

The airframe production complex under investigation here is assumed to be based on four relatively distinct stages of activity: raw material production, parts production, sub-assembly, and final assembly. In this study, experimental data on all four stages were gathered, but our focus here will be primarily upon the final assembly stage. This stage relies upon 20 production-assembly areas with up to 20 production-assembly points in each area, of which it is assumed that no more than 5 will be evading the agreement (producing proscribed airplanes or rockets) at any one time. Further, we postulate that when such points are engaged in evasion they will do so during a specific time period of *h* hours or *d* days, with a clear break between legal and illegal activity. Finally, since we are not considering here the inspection that would occur along transport routes or at junctions between production areas, the physical locations and distance between areas are ignored.

As to inspection procedures, we can assume that site visits (which may well, of course, be supplemented by permanent on-site inspectors, records examination, and personnel interviewing) may occur regularly every *d*th day or on the *average* of every *d*th day. Naturally the latter (stochastic or probabilistic) schedule is preferable to the deterministic one, and by using a Monte Carlo method, a random number can be generated for each site on each day to ascertain whether that site is to be visited on that day.[6] This is equivalent to tossing a well-balanced *d*-sided coin, of which all sides (days) but one are marked 'no visit'. Another postulated inspection characteristic is that the probability of correctly distinguishing between legal and illegal activity *increases*

as we move from the construction of sub-units and components through various stages of sub-assembly to the final, assembled product. At stages 1 through 3 we assume that the probability of a correct or false appraisal does not change from hour to hour or day to day; these points produce items which always look alike.

But at stage 4, where final assembly of the various components occurs, it makes a considerable difference whether the inspection team is there at the beginning or at the end of the stage, with the probability of an accurate appraisal rising steadily throughout that final stage. An the probability of detection should never reach certainty (i.e., 1.0) but when other variables are brought to bear, such a probability is nearly attainable. Furthermore, when all other variables are considered, we find that probability of detection remains at zero beyond maximum range, rises slowly as that threshold is crossed, then climbs rapidly during middle ranges, to level off again at very close ranges. Thus when the total complex (object, range, weather, crew, equipment, etc.) is considered, we get a probability-of-detection curve which looks more like the broken line AB'C shown in Figure 1.
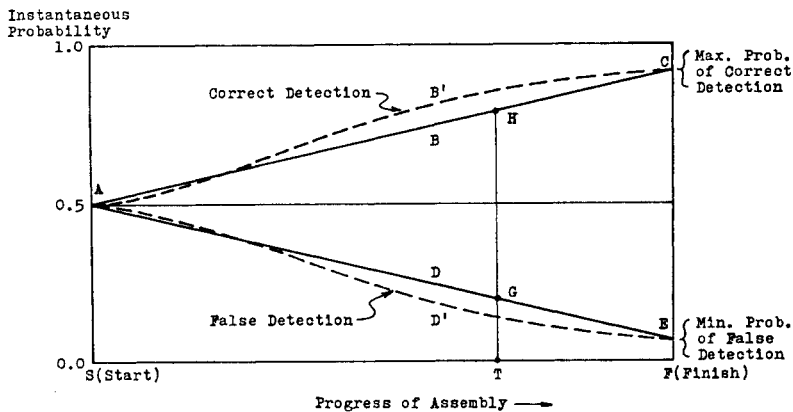

Fig. 1. *Probabilities of correct and false detection as final assembly stage progresses*

analogy from studies of radar detection may be helpful here. If we ignore for the moment the efficiency of a particular radar crew, its gear, and its search strategy, the probability of detection of an incoming aircraft will vary inversely with the range to the target.[7]

But when the constant factor of decreasing range is combined with the other variables which affect detection, the curve takes on a more complicated shape. Whereas, in theory, there is no range at which detection could not occur — and the curve therefore never reaches the zero level — in practice there is such a range. Likewise, again considering range alone, The same situation may be assumed to obtain in the matter of detecting evasion of a weapon production ban. In the latter case, time takes the place of distance, but with the time conceived of in terms of proximity to completion rather than proximity to beginning. Moreover, our probabilities here must be expanded to include that of both correct detection (when evasion *is* taking place) and false detection (when evasion is *not* taking place). The precise nature of the curve which would illustrate the relationship between progress of the final assembly stage and the probabilities of both correct and false detection is far from certain.

It *could* follow the inverted J-curve or the flatter S-curve. But for our purposes the shapes of these curves are less important than the direction, so we will settle for a linear presentation, with the probability of correct detection rising over time as shown by line ABC in Figure 1, and that of false detection falling as shown by line ADE. Suppose the assembly process takes *SF* days to complete and that the inspection method assigned to this process has a probability of ultimate correct detection *FC* and a probability of false detection *FE*. Then, on the *T*th day the probability of correct detection of an evasion is *TH* and the probability of a false detection (if the work is indeed non-evasive) is *TG*. At the start of stage 4, the probability of either correct or false detection by an unbiased inspectorate is .5, since *ST* equals zero. And at the end of this final assembly stage, with *ST* equal to *SF*, our probabilities are *FC* and *FE* respectively, which are the full values expected of the particular inspection method, under normal conditions. Instead of using the graph, however, these values may be computed by use of the following equations:

TH (prob. of correct detection at time T)

$$= .5 + (FC - .5)\ \frac{ST}{SF}$$

TG (prob. of false detection at time T)

$$= .5 - (.5 - FE)\ \frac{ST}{SF}$$

Later on in our report, such calculations will be presented.

3. *The simulation procedure*

Turning now from a description of the inspection simulation situation to the actual procedures employed, the first point to emphasize is that this is essentially a semi-dynamic model. If a fully dynamic model were being utilized, we could actually modify the values of most or all of the inputs *during* any particular simu-lated run; that is, the program would actually pick up, process, and adapt to, any new data turned up during the run. A static model, on the other hand, does little more than perform a large number of running pre-programmed computational assignments. The model used here falls somewhere between the static and the dynamic. For example, no provision is made for the inspectors rearranging their schedules and techniques when certain suspicion thresholds are crossed, or for the evader altering fundamentally his production scheme as he ascertains the nature of inspection methods. On the other hand, even though the inspection procedure or evasion scheme, once instituted, do not undergo change during a given run, the evasion scheme need not (and does not) necessarily commence and terminate simultaneously with the beginning and end of the run. Nor are there any inhibitions on the freedom of both inspector and evader to modify from one run to the next. Moreover, the Monte Carlo procedure, by producing randomized inspection results over time, simulates a degree of interaction between the two. In practice, of course, there must be a fair degree of such learning and revising, but the non-adaptive features used here do permit us to compare the relative effectiveness of each inspection method under varying production systems and evasion schemes. Had we included the adaptive features of a fully dynamic model, the print-out results would have been largely a welter of data emerging out of uncontrolled variables. The alternative is either a much more fully articulated decision model or the introduction of human players, with all of the attendant costs of the former and the weaknesses of the latter.[8] Our reasoning is thus reflected in the schematic flow diagram of Figure 2 with the computer programmed either to retain the original evasion scheme and inspection system at the beginning of each new run or to introduce new ones.
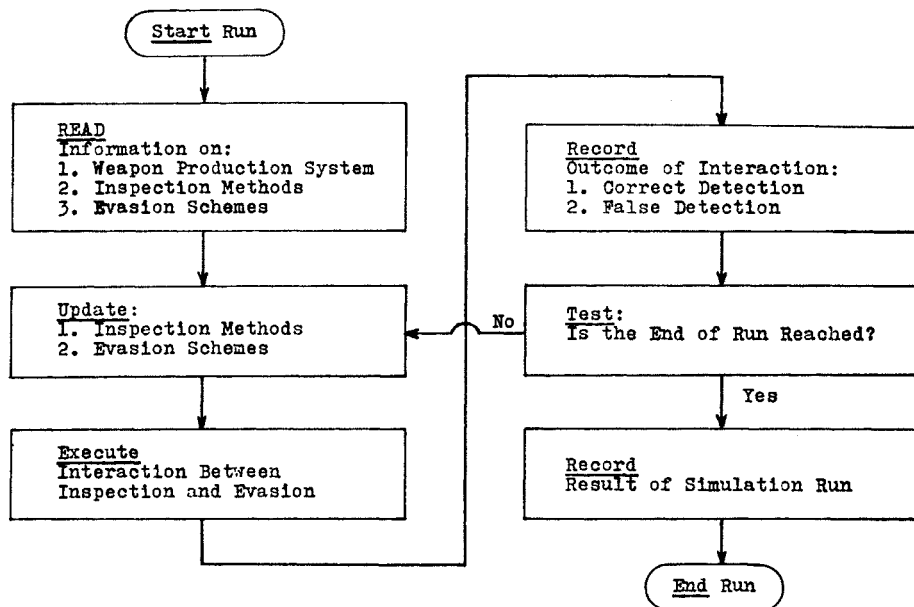
Fig. 2. *Schematic flow of simulated inspection*

The Simulation Inputs

The next operation in designing the computer simulation is that of identifying each of the variables that will affect the outcome of any particular inspection run. There are three sets of variables: (a) the production system, (b) the evasion scheme employed by the production system, and (c) the method used by the inspection team in the particular run. In greater detail:

(A) Production System
1. How many of the 20 production areas are operating in each of the 4 production stages
2. How many of the 20 production points are operating in each production area
(B) Evasion Scheme
1. No. and identity of production areas evading in each stage
2. No. and identity of evading production points in each of above areas
3. Starting and finishing days of evasion in each evading stage
(C) Inspection Method
1. Average probability of correct detection assumed for each run in every stage

2. Average probability of false detection assumed for each run in every stage
3. Frequency and time pattern of on-site inspections for each production area

A note regarding the values assigned to variables C-1 and C-2 is in order here. From general experience in a wide variety of settings (industrial quality control, medical diagnosis, law enforcement, pattern recognition, etc.), there are certain well-established performance ranges for specific operations under 'normal' conditions.[9] On the basis of these, we assign to each of our three inspection methods the following *normal* expected performance characteristics:

	A	B	C
Probability of Correct Detection	.70	.80	.90
Probability of False Detection	.04	.08	.12

These rough probabilities are, however, based on hypothetical long-run conditions of normality, and fail to take account of momentary perturbations in the production system, the inspection methods, and the environment within which any par-

ticular inspection occurs. In order to introduce greater realism into our simulation, then, we employ a Monte Carlo technique by which the computer generates from a set of equi-probable numbers, such random numbers as to reflect the inevitable short-run fluctuations in the inspection process; they are used to modify (up or down) the assumed normal probabilities of correct and false detections in any given run.

The Inspection Process

Once the various inputs have been decided upon, and appropriate values assigned or calculated, the program proceeds as follows. (See Figure 2 and Figure 5 for the simplified and the detailed flow diagrams, respectively.) Using an IBM 7090 and a program written in MAD (Michigan Algorithm Decoder) with approximately 1,000 contingent instructions, we execute the following steps. For each run, values are assigned as above, with the computer instructed as to which production system is being utilized, where and when evasions are *actually* occurring in that production system, and which inspection techniques to employ. For each of the 400 iterations, we modified each set of inputs, so that almost all conceivable combinations of conditions could be experienced, but as Tables 1 and 2 indicate, only a small sample of our data is reported here.

During the simulation runs covering stages 1 through 3, the following calculations were made. For each *evading* production point, a random number (R_1), in the interval 0 to 1, is generated by the computer and then compared to the modification of the probability of correct detection (P_1) inherent in the particular inspection method used in that run. If P_1 turns out to be greater than or equal to R_1, the evasion has been correctly detected, but if P_1 is less than the generated value R_1, the evasion has gone undetected.[10] For each *non-evading* production point, another

random number (R_2) in the interval 0 to 1 is generated and then compared, as in the evading situation, with the modified probability of false detection (P_2) applicable to the particular inspection method utilized in that inspection run. If R_2 turns out to exceed P_2, the run is said to have correctly observed that no evasion was occurring, but if R_2 is equal to or less than P_2, a false detection is recorded.

In stage 4, as suggested earlier, the computations are more complex. Here we have to account for the number of days which this final assembly stage requires and for the number still remaining when the inspection occurs. Using the equation discussed earlier (in connection with Fig. 1) we calculate P_3, the modified probability of correct detection at the end of day T of a production point which is evading:

$$P_3 = .5 + \frac{(P_c - .5)\,(T - T_s + 1)}{(T_f - T_s + 1)}$$

where

P_3 — modified probability of correct detection on inspection date (time T)

T — present date (of inspection visit)

P_c — probability of correct detection at time T_f

T_f — date on which this assembly finishes

T_s — date on which this assembly starts.

Once P_3 is calculated, a third random number (R_3) in the interval 0 to 1 is generated, and the two are compared. If P_3 turns out to be equal to or greater than R_3, the evasion has been correctly detected, but if P_3 is less than R_3, it has gone undetected.

For a *non*-evading production point in stage 4, we first calculate P_4, the modified probability of a false detection at the end of day T:

$$P_4 = .5 - \frac{(.5 - P_f)\,(T - T_s + 1)}{(T_f - T_s + 1)}$$

where

P_4 — modified probability of false detection on inspection date (time T)

T — present date (of inspection visit)

P_f — probability of false detection at time T_f

T_f — date on which this assembly finishes

T_s — date on which this assembly starts. Once P_4 is calculated, a random number (R_4) in the interval 0 to 1 is generated, and the two are compared. If R_4 turns out to exceed P_4, the inspection run is said to have correctly observed that no evasion was occurring, but if R_4 is equal to or less than P_4, a false detection is recorded.

The Simulation Output

As each run of the varying inputs proceeds through the simulated inspection process, and the multiple calculations are completed, the computer reads out a wide range of inspection performance results. These outputs initially take two forms:

internal information and external information. The external information is equivalent to the inspectorate's report: which areas and points in each of the first three stages are *believed* to be engaging in illegal production or assembly activities. In addition, for stage 4 (final assembly) the report indicates the point of progress between start and finish at which evasion is alleged to have occurred.

This external information report is then checked by the computer against the internal information: which areas and points had, indeed, been engaged in illegal activity. In evaluating the inspectorate's report, it prints out the distribution of correct detections where evasion had in fact been taking place, and false detections for all the non-evading, but suspected, areas and points. At the close

Table 1. *Percentage of correct and false detections in exact numbers : five evasion schemes vs. three inspection methods during 400 inspection visits*

Percentage Frequency of Correct Detection					Percentage Frequency of False Detection				
No. of Evading Points	Exact No. of Correctly-detected Evading Points	Methods			No. of Evading Points	Exact No. of Falsely-detected Non-evading Points	Methods		
		A	B	C			A	B	C
1	0	31.5	19.2	10.8	4	0	97.2	92.7	86.0
	1	68.5	80.8	89.2		1	2.8	7.3	14.0
	0	8.0	5.2	0.0		0	92.5	87.2	77.0
2	1	45.8	34.5	19.5	3	1	7.2	11.5	21.2
	2	46.2	60.3	80.5		2	0.3	1.3	1.8
	0	4.2	0.8	0.5		0	87.5	76.7	71.2
3	1	19.8	8.2	2.8	2	1	11.5	22.3	26.8
	2	44.5	39.2	25.0		2	1.0	1.0	1.5
	3	31.5	51.8	71.7		3	0.0	0.0	0.5
	0	0.0	0.0	0.0		0	85.2	72.7	61.0
	1	6.8	1.3	0.8		1	14.3	24.5	32.7
4	2	27.0	17.5	6.7	1	2	0.5	2.0	5.5
	3	42.2	36.2	27.0		3	0.0	0.8	0.8
	4	24.0	45.0	65.5		4	0.0	0.0	0.0
	0	0.0	0.0	0.0		0	76.2	62.7	52.9
	1	3.2	1.0	0.0		1	21.0	31.7	36.1
5	2	15.0	5.2	1.0	0	2	2.8	5.3	9.4
	3	28.8	22.0	9.0		3	0.0	0.3	1.4
	4	35.2	40.8	29.8		4	0.0	0.0	0.2
	5	17.8	31.0	60.2		5	0.0	0.0	0.0

of a particular 400-day run, average values for correct and false detections are computed and printed out, providing the experimenters with the data by which the performance of each inspection method *vis-à-vis* each evasion scheme may be appraised.

4. *Simulation results*

In Tables 1 and 2 and Figures 3 and 4, we present a small part of the output generated in the 18 experiments of 400 days each — or 7,200 simulated daily inspection visits representing about 3 hours of IBM 7090 computer time. The Table indicates the performance of our three different inspection methods when pitted against five different evasion schemes (classified by number of production points evading during the run). On the left-hand side of Table 1, we see the frequencies with which each inspection method correctly detected precisely the number of points listed when the number actually evading ranged from 1 to 5. To illustrate, when 4 points were evading, method A detected 3 (and only 3) of them 42.2 per cent of the time and all 4 of them only 24 per cent of the time; and when 5 were evading, method C detected exactly 4 of them only 29.8 per cent of the time, but detected all 5 on 60.2 per cent of the occasions. These data are not, however, of much interest to the designers of an inspection system as they stand. Their usefulness only becomes apparent when we look at threshold results, i.e. the *minimum* number of correct detections or the *maximum* number of false detections. In other words, we do not care much if method A, when con-

Table 2. *Percentage of correct and false detections above and below various thresholds : five evasion schemes vs. three inspection methods during 400 inspection visits*

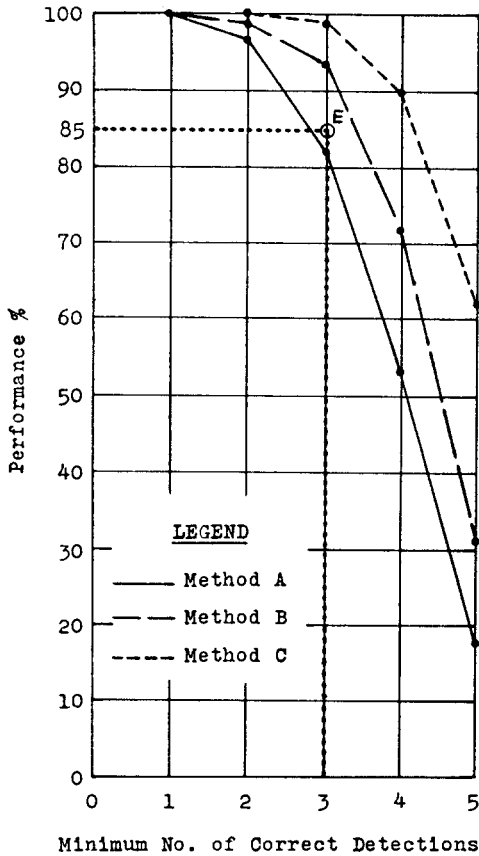| Percentage Frequency of Correct Detection					Percentage Frequency of False Detection				
No. of Evading Points	Minimum No. of Correctly-detected Evading Points	Methods			No. of Evading Points	Maximum No. of Falsely-detected Non-evading Points	Methods		
		A	B	C			A	B	C
1	1	68.5	80.8	89.2	4	0	97.2	92.7	86.0
						1	100.0	100.0	100.0
2	1	92.0	94.8	100.0	3	0	92.5	87.2	77.0
	2 (both)	46.2	60.3	80.5		1	99.7	98.7	98.2
						2	100.0	100.0	100.0
3	1	95.8	99.2	99.5	2	0	87.5	76.7	71.2
	2	76.0	91.0	96.7		1	99.0	99.0	98.0
	3 (all)	31.5	51.8	71.7		2	100.0	100.0	99.5
						3	100.0	100.0	100.0
4	1	100.0	100.0	100.0	1	0	85.2	72.7	61.0
	2	93.2	98.7	99.2		1	99.5	97.2	93.7
	3	66.2	81.2	92.5		2	100.0	99.2	99.2
	4 (all)	24.0	45.0	65.5		3	100.0	100.0	100.0
						4	100.0	100.0	100.0
5	1	100.0	100.0	100.0	0	0	76.2	62.7	52.9
	2	96.8	99.0	100.0		1	97.2	94.4	89.0
	3	81.8	93.8	99.0		2	100.0	99.7	98.4
	4	53.0	71.8	90.0		3	100.0	100.0	99.8
	5 (all)	17.8	31.0	60.2		4	100.0	100.0	100.0
						5	100.0	100.0	100.0

Fig. 3. *Frequency distribution curves for correct detection – five evading points inspected by three methods*
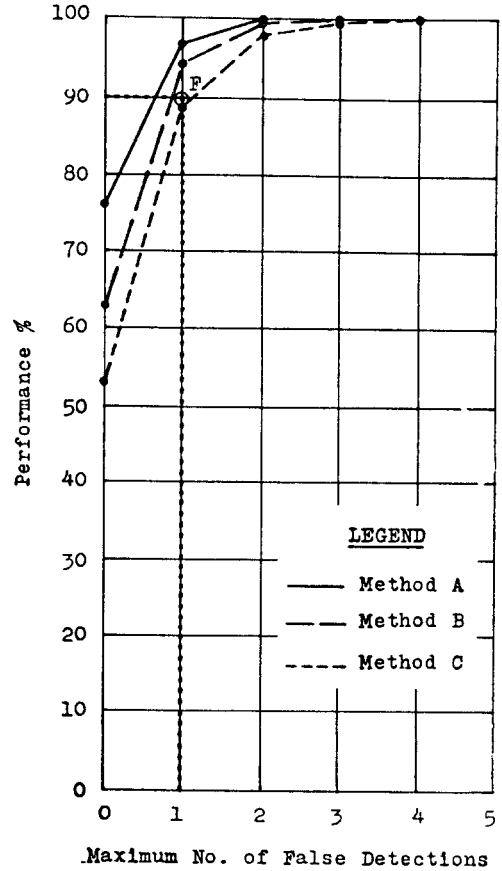

Fig. 4. *Frequency distribution curves for false detection – five non-evading points inspected by three methods*

fronted with five evading points (to take one of the examples illustrated on the left-hand side of Table 1) detects 4, and only 4, of the 5 points with a frequency of 35.2 per cent. But we do care that this inspection method in this evasion case detects 4 *or more* (4 or 5) of the 5 points with a frequency of 35.2 + 17.8, or 53.0 per cent of the time, as shown in Table 2. Thus, in order to ascertain the simulated performance of a given method, we add together all the specific frequencies of correct detection (shown in Table 1) at and above the performance we desire, in order to obtain the results shown in Table 2 and Figures 3 and 4. The lower the performance we can accept in designing an inspection method, the further up the col-

umn we go and the higher we find the frequency of correct detection *of that number or more* to be. Using this same illustration (Table 2 and Figure 3) once again, we note that method C detected correctly all 5 of the 5 evading points on 60.2 per cent of the occasions, detected 4 *or* 5 of them 90.0 per cent of the time, and 3 or 4 or 5 of them 99.0 per cent of the time.

Turning to the equally important matter of performance in *avoiding false detections*, we apply the same procedure, but in reverse. That is, when method B was confronted with 3 evading points in a given production area, it did not make a single false detection on 87.2 of its inspection visits, and it made 1 or less (no more than 1) on 98.7 of its iterations.

5. Evaluating inspection methods

Having summarized the results of a small portion of our 7,200 simulated inspections, let us now look somewhat more closely at the relative efficacy of the three methods employed, bearing in mind the significance not only of correct detection, but of false detection and the general political costs of overintrusiveness. For convenience, we summarize their *normal* expected performance characteristics again:

	A	B	C
Probability of correct detection	.70	.80	.90
Probability of false detection	.04	.08	.12

In Table 2, covering a typical 400-day run of inspection and evasion activity, certain sharp distinctions are evident. For example, when only 1 of the production points was engaged in illegal activity, method C did better (89.2 per cent detections) than either A (68.5) or B (80.8) in correctly detecting that evasion, as might be expected from its normal capabilities; but, also as expected, it revealed the lowest capacity to avoid false alarms, reporting no false detections in only 61.0 per cent of the time, whereas A and B completely avoided false detections 85.2 and 72.7 per cent of the time, respectively. Or, when 4 points were evading, method C surpassed A and B in detecting all of them and in detecting any portion of them, but also exceeded them in false detection frequencies. And when 2 production points were evading (and 3 were therefore *not* evading), methods A and B never generated (always avoided) more than 1 false detection whereas method C occasionally produced 2 false detections. In general, then, because of the greater capability (reflecting equipment, personnel, training, access, etc.) designed into method C, its average performance is better in detection but worse in terms of false detection rates, as clearly evidenced in Table 2.

This generalization leads, then, to a final concern for inspection system design: what trade-offs are possible between high detection rate and low false-alarm rate? Bearing in mind the oft-repeated reminder that a 'foolproof' inspection system may be either technologically impossible to build or so expensive in fiscal or in political terms as to be unnegotiable, we might look briefly at the relative efficiency of these three methods at the sub-perfect levels.[11] For example (Table 2), if we were willing to accept a threshold of detection which correctly identifies 1 out of 2 evading points, methods A and B might begin to look more attractive. That is, the discrepancy between B's and C's performance is considerably lower at the 1-out-of-2 level than it is at the 2-out-of-2 level, and in return for that remaining (but diminished) inferiority, we get 100 per cent ability to avoid any crossing of the 2-or-less false alarm threshold, a slightly higher ability to stay at the 1-or-under level, and an even greater ability to avoid false alarms entirely. Moving up to a more ambitious evasion scheme, with 3 of the points in a given production area engaged in illegal activity, we find that methods B and C detected at least 1 of the evasions more than 99 per cent of the time, that B detected 2 or more on 91 per cent of its occasions (compared to 96.7 per cent for C) and that A produced fewer false alarms than C or B.

Methods similar to A and B have much to commend them. If an inspection system is realistically designed, it will avoid two dangers. First, it will not convert a single detected violation into a major crisis for the signatory governments or the inspection or enforcement agencies, but will recognize the large number of factors that can produce either an unintended violation or a false alarm; cumulative evidence rather than isolated clues must provide the basis for response. Second, and in the same vein, it is by no means necessary to 'get *all* the goods' on an intentional violator; cumulative evidence of continued and widespread violation should certainly suffice, even though there may have been

more violations than were detected. Given these two considerations alone, it seems perfectly reasonable to design and institute an inspection arrangement which reduces the probability of dangerous false alarms and at the same time offers a high probability (not a certainty) of detecting *most* (not all) violations.

We shall now turn more systematically to the problem of evaluating inspection methods. Using the data shown in Tables 1 and 2, we suggest here two possible, and relatively simple, ways of calculating the relative efficacy of such alternative methods as may be available. One is a performance-only approach and the other is a cost-computation approach.

6. *A performance-only selection*

Perhaps the most simple means of arriving at a quick, and necessarily tentative, decision as to which inspection method to use is that of selecting more or less intuitively acceptable minimum performance thresholds for both correct detection and false detection, and then ascertaining which methods meet these requirements under experimental conditions. In Figures 3 and 4 we have plotted frequency polygons from the data in Table 2. Each point on the polygons shown in Figure 3 represents the frequency with which that particular inspection method *correctly* detected (out of the five evading units) *at least* the number of units shown on the horizontal axis. Likewise, each point on Figure 4, which represents the inspection of 5 non-evading units, indicates the frequency of *falsely* detecting the *maximum* number of units shown on the horizontal axis.

Since we have here data with 5 evading and 5 non-evading units — a total of ten units — we shall illustrate a case for a performance-only selection involving a plant producing 10 units or airframes daily, of which half are evading. To deal with a case with different numbers of evading and non-evading units, we merely construct the two graphs for those specific numbers, using the data from Tables 1 and 2.

Using the graphs in Figures 3 and 4, let us suppose that our minimum requirements for an inspection method are these: 1) the inspection method must *correctly* detect, 85 per cent of the time, at least 3 of the evading units when 5 of the 10 units are actually engaged in evasive activity, and 2) the number of falsely detected units among the 5 non-evading units should not exceed 1 at least 90 per cent of the time. To find the point of threshold for the first (correct detection) requirement in Figure 3 we draw a horizontal line from the 85 per cent point on the vertical (performance) axis and a vertical line up from the 3 point on the horizontal (minimum number of detections) axis and obtain a point of intersection, E. Since point E indicates the minimum acceptable performance, any inspection methods whose polygons lie either on or above this point are acceptable. A glance shows that methods B and C satisfy this requirement. In order to find the point of threshold for the second of the above criteria (false detection), we draw, on the graph in Figure 4, a horizontal line from the 90 per cent point on the vertical axis and a vertical line from the 1 point on the horizontal axis and obtain a point of intersection, F. Point F indicates the maximum acceptable level of false detection, and any inspection methods whose polygons lie either on or below this point are acceptable. We find that methods A and B meet this requirement. Thus, methods B and C meet the criterion for *correct* detection and methods A and B meet those for *false* detection. Since method B is the only one to satisfy both criteria, we select this inspection method.

One of the main weaknesses of this selection method lies in specifying, as the case for examination, only the single combination of 5 evading and 5 non-evading units, whereas we may obtain from 10 inspection units up to 11 possible

combinations of evading and non-evading units; these combinations have either no evading point or from 1 to 10 evading points. Another weakness is that it does not take into consideration any of the factors related to costs of installing and maintaining the inspection system, factors which would become crucial if more than one method satisfied the initial criteria. Thus, we now turn to a somewhat more elaborate mode of selection.

7. *Selection by cost comparison*

As we suggested above, any serious scheme for comparing and evaluating alternative inspection methods must not only consider the *performance* of each method, it must also consider such elusive factors as political acceptability and intrusiveness as well as the costs of equipment, personnel, and maintenance.[12] In this section, employing the central notion of efficiency, we propose a technique by which the latter of these considerations may be meaningfully combined with performance by using a single measuring unit.

Basically, we have two logical choices when confronted with the need to select an inspection method. Either (a) we can specify the minimum acceptable performance level (as in the previous section) and then ascertain which of the available methods meeting or exceeding that level is least costly, or (b) we can specify the maximum monetary cost we are willing to sustain and then ascertain which of the methods remaining below that ceiling is likely to give us the highest performance in terms of correct and false detection. Each individual approach has its merits, but here we seek to combine the two in such a fashion as to express both criteria by use of a single index. Our first step is to postulate a range of alternative evasion schemes, each characterized by the probability of a given combination of evading and non-evading units; the identity of these units is not, however, material.

Using these probabilities, the performance values of each inspection method (such as shown in Table 1) and costs, we may compute the expected total cost of the method through a formula of mathematical expectation. Given this selection strategy, our next problem is to develop a procedure by which such readily estimable material costs as initial installation and daily maintenance of the inspection system may be *combined* with a variety of non-material, political and psychological costs into a quantitative index; for this index we will use a hypothetical 'dollar'. Among those costs that will need to be calculated in such dollar terms are (a) the failure to detect actual evasion, and (b) the registering of false detections. For example, if the inspection system were to fail in detecting evasion at certain points or stages, such failure must be compensated for at other stages or points in order to remain politically acceptable. Such 'beefing up' of the system at these points costs money. Or if one signatory were not confident of the system's performance at one production stage or another, it might well be expected to increase its expenditures for certain non-proscribed weapons (including perhaps passive defense measures) as a matter of insurance against possible military disadvantage.

On the other hand, false detections are likely to create resentments on the part of the offended signatory, producing a drop in his cooperation with the inspectorate. As obstacles and roadblocks are put in the path of the inspection system, a greater investment will be necessary in order to restore the originally-sought performance levels and to compensate for the inevitable sabotage of the system. At the extreme, a continuous barrage of false detections could lead to unilateral abrogation of the agreement, the costs of which cannot fail to be impressive. Despite the difficulty of converting these considerations into our single monetary index, it seems to be worth the effort.

8. Formulation of inspection costs

In the previous section, we proposed considering the total cost of an inspection method as composed of four variables: initial installation, maintenance and operation, failure to detect evasion (i.e. successful evasion), and false detection. Let us discuss each briefly in turn.

The initial cost (CINI) represents the original, fixed investment in surveillance equipment, buildings, transport facilities, data processing machinery, clerical materials, and basic personnel. We assume that none of this investment can be recovered at the end of an inspection period. For maintaining and operating the system, regular expenditure on material and personnel is essential. This operating cost (COP) is assumed to be a linear function of the number of days, M, constituting the inspection period; thus total operating cost (TCOP) is:

$$TCOP = COP \cdot M \qquad (1)$$

If we denote the total cost of successful evasion by TCSUCEVA and that of false detection by TCFALDET, the total cost of the entire inspection system, being denoted by TC, is:

$$TC =$$
$$CINI + TCOP + TCSUCEVA + TCFALDET \qquad (2)$$

In the following section, we shall elaborate these mal-performance costs of TCSUCEVA and TCFALDET, assuming that these costs are directly proportional to the numbers of successful evasions and false detections, respectively.[13]

The effectiveness of an inspection system is governed by the specific conditions under which it operates: the total number of production units per day (assumed to be constant) and the number among this total that are illegal. If we assume a production total of N per day, then there are $N+1$ possible evasion conditions, ranging from zero through N. For each of these conditions, we assume a probability of evasion

(PEVA$_i$) where i is the number of illegal units $(i = 0, 1, 2, ..., N)$. Since there are M days in a given inspection period, the expected number of days (D) in which the condition with i illegal units occurs can be represented by the following:

$$D_i = M \cdot PEVA_i \quad (i = 0, 1, ..., N) \qquad (3)$$

Next, we determine the effectiveness of an inspection system on a day when i units among the total of N units are evading. In this case, the system might not detect any evading units or it might detect anywhere from one to as many as all of the evading units. Let us use PDET$_{i,j}$ to denote the probability of detecting exactly j units $(j = 1, 2, ..., i)$ out of i evading units by this inspection system. Then, the expected number of evading units correctly detected under this condition (denoted by CORDET$_i$) is the following sum:

$$CORDET_i = 1 \cdot PDET_{i,1} + 2 \cdot PDET_{i,2} +$$
$$+ 3 \cdot PDET_{i,3} + \cdots + i \cdot PDET_{i,i}$$
$$= \sum_{j=1}^{i} j \cdot PDET_{i,j} \quad (i = 1, 2, ..., N) \qquad (4)$$

The expected number of units which have successfully evaded among the total of all i illegal units (SUCEVA$_i$) is obtained by subtracting the above number, CORDET$_i$, from the total number of evading units:

$$SUCEVA_i =$$
$$i - \sum_{j=1}^{i} j \cdot PDET_{i,j} \quad (i = 1, 2, ..., N) \qquad (5)$$

Since the expected number of days having i evasions during the period is D_i and the expected number of the illegal units that have successfully evaded in any of these days is SUCEVA$_i$, the total expected number of successful evasions during the entire period is the product $D_i \cdot SUCEVA_i$ or $M \cdot PEVA_i \cdot SUCEVA_i$. Therefore, if CSUCEVA represents the cost of having one successful evasion, then the total cost of successful evasion during the

period is the following, combining the costs for all i $(i=1, 2, ..., N)$:

$$\text{TCSUCEVA} =$$
$$\text{CSUCEVA} \cdot M \sum_{i=1}^{N} \text{PEVA}_i \left(i - \sum_{j=1}^{i} j \cdot \text{PDET}_{i,j}\right) \quad (6)$$

Let us now turn to a calculation of the total cost of false detection. If $\text{PFALDET}_{i,j}$ denotes the probability of falsely detecting exactly j units $(j=1, ..., i)$ out of the total of i legal units, then the expected number of falsely detected units is given by the following:

$$\text{FALDET}_i = 1 \cdot \text{PFALDET}_{i,1} +$$
$$2 \cdot \text{PFALDET}_{i,2} + \cdots + i \cdot \text{PFALDET}_{i,i} =$$
$$\sum_{j=1}^{i} j \cdot \text{PFALDET}_{i,j} \quad (i=1, 2, ..., N) \quad (7)$$

Since PEVA_i represents the probability of having exactly i evading units among the total of N points observed a day, the probability of having exactly i legal (non-evading) units is PEVA_{N-i}. Therefore, the product of $M \cdot \text{PEVA}_{N-i}$ represents the expected number of days in the period in each of which exactly i legal units occur. If we let CFALDET be the cost of each false detection, then following the same method used to obtain TCSUCEVA, we may write the total cost of false detection, being denoted by TCFALDET, as follows:

$$\text{TCFALDET} = \text{CFALDET}$$
$$M \sum_{i=1}^{N} \text{PEVA}_{N-i} \sum_{j=1}^{i} j \cdot \text{PFALDET}_{i,j} \quad (8)$$

Substituting the right-hand side of Eq. (1) and those of Eqs. (6) and (8) into TCOP, TCSUCEVA and TCFALDET of Eq. (2),

$$TC = \text{CINI} + \text{COP} \cdot M + \text{CSUCEVA}$$
$$M \sum_{i=1}^{N} \text{PEVA}_i \left(i - \sum_{j=1}^{i} j \cdot \text{PDET}_{i,j}\right)$$
$$+ \text{CFALDET} \cdot M \sum_{i=1}^{N} \text{PEVA}_{N-i} \sum_{j=1}^{i} j \cdot \text{PFALDET}_{i,j} \quad (9)$$

9. *Illustrations*

In attempting to illustrate the cost effectiveness approach developed here, we have at least two options. In one case, we can assume that: (a) within any day's production, all units in either of the evading and non-evading groups are homogeneous, i.e. have an identical appearance, and (b) the size of the group has no effect on the performance of the selected inspection method. In this case, we can treat PCORDET and PFALDET as binomial distributions, and in the Appendix the appropriate calculations for FALDET and SUCEVA are presented. In the second option, and the one selected here, we seek greater realism; i.e. assume lack of homogeneity and assume that the size of either of the evading and non-evading groups does indeed affect performance. Let us turn, then, to the illustration, in which the three inspection methods utilized in the original simulation are evaluated. It will be recalled that a total of five units per day were being produced and inspected, that anywhere from zero to five might be evading, and that the performance data in Table 1 show the frequencies with which each inspection method detected each specific possible number of evasions. Further, we postulate the following (purely hypothetical) initial costs and daily operating costs for a typical airframe installation:

	Method A	Method B	Method C
Initial Cost (CINI)	$1,200,000	$900,000	$1,500,000
Daily Operating Cost (COP)	$4,000	$7,000	$5,000

respectively, we obtain the following final form of the total cost of our inspection system:

For this illustration, we use an inspection period made up of 360 days. Our two hypothetical evasion schemes are as

follows: in scheme I, there are 5 phases of 72 days each; beginning with only 1 evasion in the first phase, this scheme adds an additional evasion at the beginning of each phase, so that there are 5 illegal units during the final (5th) 72-day phase. In scheme II, there are 4 phases of 90 days each; in the first phase there are 3 illegal units, in the second phase there are 2, in the third phase 1, and in the final (fourth) phase there are none. Of course, the sequence in which these evasions occur is of no consequence; what does matter is the distribution of daily evasion rates. Moreover, the evader would be unlikely to follow so simple a scheme and would probably scatter his evasion phases into discontinuous periods. Our calculation assumes such discontinuity of phases, but for purposes of clarity we present the evasion schemes as follows:

but eliminate the multiplier, CSUCEVA, which is the *cost* of a successful evasion:

Total Expected No. of Successful Evasions

$$= M \sum_{i=1}^{5} \text{PEVA}_i(i - \sum_{j=1}^{i} j \cdot \text{PDET}_{i,j})$$

$$= 360 \times \frac{1}{5}\{(1 - 1 \times .685 + 2 - (1 \times$$

$$.458 + 2 \times .462) + 3 - (1 \times .198 +$$
$$2 \times .445 + 3 \times .315) + 4 - (1 \times 0.68 +$$
$$2 \times .270 + 3 \times .422 + 4 \times .240) + 5 -$$
$$(1 \times .032 + 2 \times .150 + 3 \times .288 +$$
$$4 \times .352 + 5 \times .178)\} = 329.18$$

Using Eq. (8) in a similar manner as we did with Eq. (6), we now obtain the expected number of falsely-detected non-evading units for method A:

Total Expected No. of False Detections

	Evasion Scheme I			Evasion Scheme II		
	No. of days	No. of evading units	No. of non-evading units	No. of days	No. of evading units	No. of non-evading units
Phase 1	72	1	4	90	3	2
Phase 2	72	2	3	90	2	3
Phase 3	72	3	2	90	1	4
Phase 4	72	4	1	90	0	5
Phase 5	72	5	0			

Postulating that the policy-makers have estimated that the dollar equivalent of failing to detect (i.e. successful evasion) is $20,000 per successful evasion and that of false detections is $30,000 per accusation — bearing in mind the previous footnote — we may now go ahead and calculate the various costs. First, however, it is necessary to ascertain the total expected number of successful evasions and false detections during the entire period for each method. For the total expected *number* of successful evasions (employing method A in this illustration) we return to the total evasion cost equation, Eq. (6),

$$= M \sum_{i=1}^{5} \text{PEVA}_{N-1}(\sum_{j=1}^{i} j \cdot \text{PFALDET}_{i,j})$$

$$= 360 \times \frac{1}{5} (1 \times .028 + 1 \times .072 + 2 \times .003 +$$

$$1 \times .115 + 2 \times .010 + 1 \times .143 + 2 \times .005)$$
$$= 28.37$$

Likewise, we can compute the total expected number of both successful evasions and false detections for methods B and C; they are 217.15 and 55.15 for B and 112.32 and 83.66 for C. Using these six values, we may now compute the total cost of all three methods, by combining all of our values as follows:

Cost Comparison
(In thousands of dollars)

	Method A	Method B	Method C
1. Cost of each successful evasion	$20	$20	$20
2. Expected no. of successful evasions per entire period	329.18	217.15	112.32
3. Total expected cost of successful evasions (1) × (2)	$6,584	$4,343	$2,246
4. Cost of each false detection	$30	$30	$30
5. Expected no. of false detections	28.37	55.15	83.66
6. Total expected cost of false detections per entire period (4) × (5)	$851	$1,654	$2,510
7. Operating cost per day	$4	$7	$5
8. Total operating cost for 360 days	$1,440	$2,520	$1,800
9. Cost of initial installations	$3,000	$2,000	$2,600
10. Total cost of inspection (3)+(6)+(8)+(9)	$11,875	$10,517	$9,156

Examining the results of these calculations, it would seem that, if one considers the four sets of variables indicated, method C is the least costly. Though C's costs for false detection (because its performance level is lower than those of the other two on this score) is highest, that figure is more than compensated for by its lower cost on the successful evasion, initial installation, and operating dimensions. Naturally, the dollar value arbitrarily assigned to successful evasions and false detections is essentially a political and military decision, and some policy-makers might well feel that the cost of a successful evasion ought to be rated much higher, that of false detection much lower, or vice versa.

Let us now apply the same procedure to evasion scheme II, as a further illustration of this procedure. In this scheme, it will be recalled, the evader begins with 3 illegal units per day; at the end of each of the four 90-day phases in this scheme, he converts one unit back to a legal one, so that in the fourth and final phase, there

are no evasions. Employing the same computations as in evasion scheme I, we discover that the total cost of our three inspection methods are as follows: A—$8,541,000, B — $8,708,000, and C — $8,800,000.

Under this evasion scheme, method A offers the lowest cost of the three, although the differences between these three inspection methods are not as significant as those encountered when dealing with the previous scheme. The point in offering these two separate illustrations is to emphasize the need to calculate the cost of each inspection method when played off against a *variety* of evasion schemes. Moreover, in the event of such a production control agreement ever being consummated, it should be clear that evasion schemes of a far more ingenious and elaborate nature must be anticipated, if for no other reason that to reassure the domestic doubters and disarm the domestic resisters.

APPENDIX

Calculating successful evasions

Let p_c denote the probability of correct detection with regard to a single inspection unit and P_e (i,j) the probability of j successful evasions among the total of i

evasions. Then, the binomial distribution for P_e (i,j) is written as follows:

$$P_e(i,j) = \binom{i}{j}(1-p_c)^j p_c^{i-j}$$

Therefore, in this case the expected num-

Column 1 (flowchart)

(Start)

Read Input on Weapon Production System

Read Input on Inspection System

Read Input on Evasion Schemes

Set Time T = 1

(a) Advance Time by 1

Set Stage No. i = 1

(b) Advance Stage No. by 1

Set Area No. j = 1

(c) Advance Area No. by 1

Is Area j of Stage i Evading at Time T?

Yes | No (e)

Compute Cumulative Sum of Evading Points in This Stage: TEVA

Obtain Modified Prob. of Correct Detection for This Area: P₁

(d)

Column 2 (flowchart)

(d) For Stage 4, Adjust P₁ with Progress of Assembly: P₃

For Each Evading Point, Pick up a Random No. Between 0 and 1: R₁

Is Correct Detection Registered (R₁ ≦ P₁ or P₃)?

Record Correctly Detected Area No. (and Progress of Assembly)

Compute Cumulative Sum of Detected Points in This Stage: NCOR

(e) Compute Cumulative Sum of Non-Evading Points in This Stage: TNEV

Obtain Modified Prob. of False Detection for This Area: P₂

For Stage 4, Adjust P₂ with Progress of Assembly: P₄

For Each Non-Evading Point, Pick up a Random No. Between 0 and 1: R₂

(f)

Column 3 (flowchart)

(f) Is False Detection Registered (R₂ ≦ P₂ or P₄)?

Record Falsely Detected Area No. (and Progress of Assembly)

Compute Cumulative Sum of Suspected Points in This Stage: NFAL

Is This Last Area of Stage i?

Yes | No (c)

Compute Effectiveness of Inspection Method:
1. P.C.D. = NCOR/TEVA
2. P.F.D. = NFAL/TNEV

Print Results:
1. NCOR, TEVA, P.C.D., Detected Area Nos. (Assembly Progress)
2. NFAL, TNEV, P.F.D., Suspected Area Nos. (Assembly Progress)

Is This Stage 4?

Yes | No (b)

Is This End of Inspection Period?
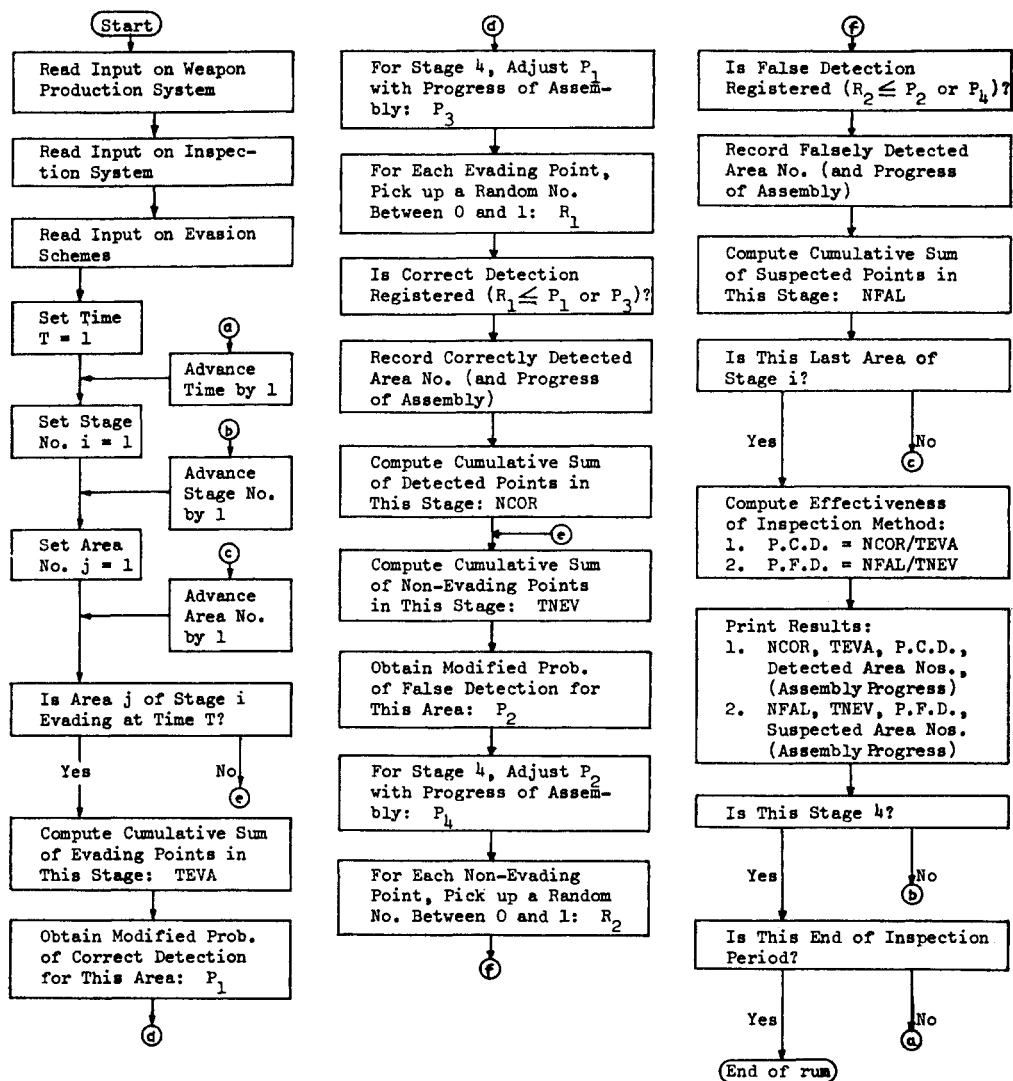
Yes | No (a)

(End of run)

Fig. 5. *Detailed descriptive flow of computer simulation*

ber of successful evasions (SUCEVA$_i$) given by Eq. (5) is replaced by the following equation:

$$\text{SUCEVA}_i = \sum_{j=1}^{i} j \binom{i}{j} (1-p_c)^j p_c^{i-j}$$

This equation is changed as follows:

$$\text{SUCEVA}_i = \sum_{j=1}^{i} \frac{j \cdot i!}{(i-j)! \, j} (1-p_c)^j p_c^{i-j}$$

$$= i(1-p_c) \sum_{j=1}^{i} \frac{(i-1)!}{(i-j)!(j-1)!} (1-p_c)^{j-1} p_c^{i-j}$$

$$= i(1-p_c) \sum_{j=1}^{i} \binom{i-1}{j-1} (1-p_c)^{j-1} p_c^{i-j}$$

$$= i(1-p_c)\{(1-p_c)+p_c\}^{i-1}$$

Thus, we have:

$$\text{SUCEVA}_i = i(1-p_c)$$

Calculating false detections

Likewise, if P_f denotes the probability of false detection with regard to a single unit, the binomial distribution for the proba-

bility of j false detections out of i non-evading units, being denoted by $P_f(i,j)$, is as follows:

$$P_f(i,j) = \binom{j}{i} p_f{}^j (1 - p_f)^{i-j}$$

And the expected number of falsely detected points (FALDET$_i$) in this case is

$$\text{FALDET}_i = \sum_{j=1}^{i} j \binom{i}{j} p_f{}^j (1 - p_f)^{i-j}$$

Following the method used for obtaining the final form of SUCEVA$_i$, the above equation can be changed to

$$\text{FALDET}_i = i p_f$$

Thus, when the probabilities of correct and false detections are binomially distributed, SUCEVA$_i$ and FALDET$_i$ are represented by very simple forms using i, p_c and p_f.

NOTES

* The authors are heavily indebted to Marcia Feingold for her technical advice in this project, and to the University of Michigan Computing Center for making its facilities available.

[1] Among the more detailed studies of the negotiations are: Bernard G. Bechhoefer, *Postwar Negotiations for Arms Control* (Washington, D.C.: Brookings Institution, 1961), Lloyd Jensen, *The Postwar Disarmament Negotiations: A Study in American-Soviet Bargaining Behavior* (Ann Arbor, Michigan: Center for Research on Conflict Resolution, 1962), and John Spanier and Joseph Nogee, *The Politics of Disarmament* (New York: Praeger, 1962).

[2] Bernard T. Feld *et al.*, *The Technical Problems of Arms Control* (New York: Institute for International Order, 1959), Louis Henkin, *Arms Control and Inspection in American Law* (New York: Columbia University Press, 1958), and Seymour Melman (Ed.), *Inspection for Disarmament* (New York: Columbia University Press, 1958).

[3] One very suggestive man-machine simulation is Robert H. Davis *et al.*, *A Game for Studying the Problem of Arms Control* (Santa Monica: System Development Corp., May 1962). Another paper which, while not experimental, deals specifically with the matter at hand is John B. Phelps, 'Some Problems of Missile-Production Inspection,' in David H. Frisch (Ed.) *Arms Reduction: Program and Issues* (New York: Twentieth Century Fund, 1961).

[4] Naturally, we are not assuming that all new production of a given weapon type will be prohibited at a given date. Rather, the assumption is that production rates will be expected to diminish according to a phased schedule, more or less coordinated with the scrapping, conversion, or transfer of existing items as well as with the build-up of 'peace-keeping machinery'. Thus the experiment reported here is primarily relevant to the later stages of any conceivable arms reduction schedule. For a detailed discussion of this phasing problem, see J. David Singer, 'Inspection and Protection in Arms Reduction', *Journal of Arms Control*, Vol. 1, No. 1 (January 1963), pp. 56–72.

[5] Some might well argue that Type I errors of commission (false detection) are not nearly as important as Type II errors of omission (successful evasion) and that the price of arms reduction is an inspection system of such rigor that false alarms of some frequency must be accepted. In the selection procedures outlined below, the decision-makers could readily build such a bias into their computations, to the extent deemed necessary. A useful discussion of the statistical implications is in Abraham Wald, *Statistical Decision Functions* (New York: John Wiley, 1950).

[6] 'Monte Carlo' is synonymous with any use of random sampling in treatment of problems of either a deterministic or a probabilistic nature. See, for example, H. A. Meyer (Ed.), *Symposium on Monte Carlo Method* (New York: John Wiley and Sons, 1956).

[7] Philip M. Morse and George E. Kimball, *Methods of Operations Research* (New York: John Wiley, 1951), pp. 134–35, and B. O. Koopman, 'The Theory of Search II: Target Detection', *Operations Research*, Vol. 4, No. 5 (Oct. 1956), pp. 503–31.

[8] An excellent analysis of the strengths and weaknesses of man-machine simulation in such a context, along with some detailed procedures and experiences, is Harold Guetzkow (Ed.), *Simulation in International Relations* (Englewood Cliffs, N. J.: Prentice-Hall, 1962).

[9] See Harry H. Goode and Robert E. Machol, *System Engineering: An Introduction to the Design of Large-Scale Systems* (New York: McGraw-Hill Book Co., 1957). Indicative of the Soviet approach to analysis of production systems is V. I. Tikhomirov, *Organization and Planning of an Aircraft Production Enterprise* (Moscow; State Publishing House for the Defense Industry, 1957).

[10] The random number which is generated on each pass is thought of as reflecting the combined deviations from the normal in terms of inspection team's efficiency on that day, perturbations in the inspection environment, and changes in the inspected objects. Thus, on a given site visit, if the combined probabilistic perturbations (R_1) add up to .85, inspection methods A and B would both be considered as having failed to detect an evasion, since R_1 is greater than the normal expected performance level of either of these two methods. Method C, of course, having a .90 normal capability is not exceeded by the random demands of this particular visit and hence is thought of — if used — as being able to have correctly detected the violation.

[11] In addition to average performance there is, of course, the matter of momentary fluctuation, so that a given method's *average* performance may be satisfactory but its occasional lapses may make it unacceptable.

[12] For a discussion of several such variables, see J. David Singer, 'Inspection and Protection in Arms Reduction', *Journal of Arms Control*, I/1 (January 1963), pp. 56–72; see also the same author's *Deterrence, Arms Control, and Disarmament* (Columbus: Ohio State University Press, 1962), *passim*.

[13] It should be reiterated that our assumption of linearity in the costs of successful evasion and false detection is for purposes of clarity only. In reality, one should expect the marginal disutilities to change, thus producing costs which are *not* directly proportional to the number of either type of failure.

SUMMARY

We have tried to present an experiment in the detection of illegal weapons production, using a rather modest and simplified computer simulation. Three inspection schemes are defined by their ability to detect evasion and to avoid false alarms, and their effects are simulated using Monte Carlo methods. Because the model is not a fully dynamic one, it employs delayed rather than more or less instantaneous feedback; the full intensity of the inspection-evasion interaction is moderated. Furthermore, the model makes a number of assumptions which are justified only in so preliminary an analysis. For simulation techniques such as that presented here to play a really useful role in the design of inspection systems and exercise a practical effect on the negotiations for disarmament, this pilot project must be considerably surpassed in complexity and sophistication.

Proceeding on the basis of data gathered during the computer simulation, we outlined a sequence of procedures whereby the potential signatories to a weapons production ban agreement might be able to sort out and compare the relative costs of alternative inspection systems outlined. We have quite consciously sought to combine such disparate material and strategic-political factors as installation and operating cost on the one hand and successful evasion or false detection cost on the other. Two caveats are thus in order. First, such 'combining of utilities' may be a perfectly reasonable procedure, but we cannot afford to forget that certain of our inputs are based on *political* judgements of a most crucial sort. Second, and equally important, the calculations suggested here incorporate only a small portion of all the variables that must be considered in selecting a man-machine inspection system upon which a nation's security may well depend.

Краткое содержание

В нашей статье делается попытка представить экспериментальный способ раскрытия нелегального производства оружия, причем мы используем довольно упрощенную предполагаемую модель. Три схемы инспекции и контроля определены их пригодностью для раскрытия попыток обхода и для предотвращения возникновения ложной тревоги. Их значение обозначено предполагаемой моделью с использованием методов Монте-Карло. Так как предложенная система не отличается большой динамичностью (она не рассчитывает на немедленное подтверждение) полная эффективность взаимодействия между инспекцией