Volume Title: Issues in Defense Economics

Volume Author/Editor: Roland N. McKean, editor

Volume Publisher: NBER

Volume ISBN: 0-87014-490-1

Volume URL: http://www.nber.org/books/mcke67-1

Publication Date: 1967

Chapter Title: Increasing Returns in Military Production Functions

Chapter Author: Malcolm Hoag

Chapter URL: http://www.nber.org/chapters/c5158

Chapter pages in book: (p. 3 - 23)

PART I

Applying Economic Concepts to Defense Problems

Increasing Returns in Military Production Functions

MALCOLM W. HOAG

THE RAND CORPORATION, SANTA MONICA, CALIFORNIA

I

Juxtaposing two authoritative quotations, perhaps bizarrely, can serve to pose the central issue for this paper, and perhaps a related one as well: First, "In every battle, concentrate an absolutely superior force (two, three, four, and sometimes even five or six times the enemy's strength), encircle the enemy forces completely, strive to wipe them out thoroughly and do not let any escape from the net." [1] Second, "Approaching the problem [of allocating resources within the Department of Defense] from the first point of view—getting the most defense from a given level of resources—we work in terms of marginal rates of transformation and substitution. Approaching the problem from the second point of view—achieving a given level of defense at the least cost, which is the way Secretary McNamara prefers to look at the problem—we work in terms of marginal products and marginal costs in order to help the top decision-maker choose the appropriate level of resources." [2]

The first quotation is taken from Mao Tse-tung, but it could easily be paralleled from Western military sources. Mr. Hitch here expresses also what has come to be an orthodox view about the new defense economics, to which he and but a small staff within the Defense De-

NOTE: Any views expressed in this paper are those of the author. They should not be interpreted as reflecting the views of The RAND Corporation or the official opinion or policy of any of its governmental or private research sponsors.

[1] Mao Tse-tung, *Selected Works*, Eng. ed., Peking, 1961, Vol. V, p. 161.

[2] Charles J. Hitch, *Decision-Making for Defense*, Berkeley, 1965, p. 52.

partment contributed so much. There is no inconsistency between these quotations, but there is a contrast.

Mao's statement is typical in its concentration upon an enemy and the choice of strategy and tactics to defeat him; analytically, it is concerned with conflict situations, and numerically with totals. This statement by Mr. Hitch can be almost as profitably applied to peaceful governmental activities where, bureaucratic conflict aside, there is no enemy; analytically, it applies best to small variations of choice within a given framework, and thus, numerically, to increments. The central issue for this paper can be put in terms of this contrast: Where and why does analytic marginalism apply well to military problems, and where not and why? Consequently when must economic analysis establish a relevant resource level before applying marginal cost/effectiveness tests? Given this central focus, the related problem is almost automatically treated: How can we improve mutual understanding and communication between military professionals and the civilian analysts of military problems who, although they are for the most part not economists by training, are applying economic principles in their quantitative tests?

Both of these questions arise, above all, in determining what aggregate military "requirements" are to be, as distinct from determining which among competing alternative weapon systems or forces is to fulfill a given requirement. They arise when the scale of military operations for a particular theater or function is itself brought into question, and therefore the total budget for it.

II

Before turning to such cases, it is appropriate to acknowledge, at least briefly, how importantly economic analysis has contributed, and can contribute, to improving choices among competing alternatives to meet a stipulated requirement, and to clarifying judgments about the issues involved in choosing a "requirement." It can contribute in many ways, of which the least tangible but perhaps most important is in stretching imaginations about what the proper alternatives are. Economists are very abstract and general when they deal with such lofty concepts as social welfare functions, or production functions for that matter. What in practice are extraordinarily diverse service inputs are typically aggregated in a highly arbitrary manner, even when programming or activity analysis techniques are employed to handle far more variables than capital and labor alone. Yet if the arbitrariness is a defect, the vision of the production process as a whole, and the attempt to structure input relationships for it, are virtues.

The administrator at the plant or base level, or the engineer, will of necessity avoid this kind of arbitrariness, which is his strength, but at the understandable cost of a narrower perspective. Faced with the problem of determining how better to accomplish a given military task, such a professional may go astray in one of two opposite directions. He may try to drive to, or even beyond, technology's outer limits in order to obtain maximum performance, regardless of opportunity costs; or, on the other hand, he may apply too narrow and restrictive a cost constraint. For all his other faults, the economist is not nearly so likely to err in these directions. While he can do more, the economist has much to contribute in countering these errors by general diagnosis and specific example. Repeated examples in practice will be needed, it can safely be forecast, for the intuitive appeal of the one error is deeply rooted in tradition and emotion as well as reason, while the appeal of the other rests in ease and practicality.

To ignore or depreciate opportunity costs in choosing among military alternatives is a practice that an economist may rightly deplore, but with which he must sympathize up to a point. No counterpart in nonmilitary pursuits is likely to have as plausible a rationale. Perhaps the closest equivalent among commercial examples would be the instance where a firm could, by virtue of but a small qualitative improvement over the products of its competitors, capture a far larger share of the market and spread largely fixed costs over its expanded production. Yet where a market is so structured that such opportunities exist, a high monetary reward for qualitative improvement is likely to be apparent to managers, even if very uncertainly estimated, and they may therefore be perfectly willing to accept high costs for attempts at superior quality.

But military improvements cannot be assessed in dollars against the test of an impersonal market mechanism, but rather must somehow be tested by planning staffs within the context of hypothetical future contingencies that one hopes will never arise. Such contexts may well put great premiums upon small quality differentials. To increase the accurate range of a gun by a small amount may matter little, because virtually all relevant battle is likely to take place well within existing ranges anyway. Surveillance problems about what to shoot at may dominate the range problem. Or it may matter very much, because in many probable situations the increased range would permit one's forces to fire effectively while staying out of range of enemy fire. Which is it? How do you know? Or change the example to another commonly cited one, greater speed for aircraft that may engage in air-to-air combat. To have a small speed margin that gives you the choice to flee or pursue successfully, but that denies either choice to your enemy, may likewise be extremely important. Or it may not, because air-to-air duels that would exploit this advantage are no longer very relevant. Again, which is it?

One thing is certain: Men who have fought in wars where they saw that some particular small qualitative advantages yielded a great edge in combat are going to have a most understandable bias toward high performance, and hang the costs. Economic argument is impeccable to the effect that real resource constraints will always be present, taking their most pronounced form in war rather than in peace, to such an extent that one military claimant can gain resources only at the expense of another, and hence that opportunity costs should be measured. But this logic runs directly counter to deep military biases. Lectures on economic fundamentals will be necessary, but insufficient to overcome these biases.

Success in overcoming these biases will only be approached through continued practice: "The question is not whether three units of alternative A are preferred to three units of alternative B, whose quality per unit is smaller. The question is whether two units of A are preferred to three units of alternative B in combat, for these combinations are of equal cost." Patience will be strained on both sides as the natural answer is reiterated: "As to type, A is preferred; as to quantity, three units are required." The economist in this dialogue needs, beyond perhaps an unaccustomed patience and a willingness to dwell on recurring fundamentals rather than move on to intellectually more challenging refinements, an understanding of both the reasons for this obdurate reply and the possible good sense in a different one that questions the given scale: "Neither of these equal-cost combinations is very relevant, but here is why (1) an alternative comparison between three units of A and four and one half units of B would be relevant, and (2) why we prefer the one to the other of these equal-cost combinations."

Misunderstandings and apprehensions by the military professional can be allayed if he perceives that the economists not only are troubling him for good reason, but for equally good reason are simultaneously criticizing those weapons system designers who, by imposing cost criteria that are too narrow, penalize future military performance unduly. Finding and indicting such narrowness is easy, although it must be emphasized that the appropriate constructive correction is hard. Favored relevant examples are supplied by ordnance: an improved bomb, shell, gun, rocket, or the like. Compare one of these with its predecessor in field tests that simulate combat conditions, and it may be fairly well determined that fifty bombs, say, of the new type will accomplish what it would take one hundred of the old to do. But fifty of the new bombs may well cost far more than twice as much as one hundred of the old. In practice, of course, the cost estimate for the new ordnance would extend over an uncertain range and might be biased downward, rather than be known precisely. Comparative effectiveness would be still more uncertain and difficult to measure, because relative effectiveness could be expected to differ markedly in different contingencies, whose relative probabilities of occurrence are not known. This oversimplified discussion, however, bypasses these vital qualifications, and presumes that relative bomb costs have been determined to be 2 : 1, while ordnance effectiveness has been determined to be less than 2:1 in relevant contingencies.

Yet to divide ordnance effectiveness by cost so measured for the new and the old, and find the new to be ruled out by cost/effectiveness criteria, is obviously wrong. The military professional will know intuitively that it is wrong. If he incorrectly thinks that Department of Defense systems analysts use cost/effectiveness criteria this way, his resistance to any and all measures of opportunity costs within the military establishment will be strengthened. If, on the contrary, he sees DOD systems analysts reversing such results, his inclination to understand and use relevant opportunity cost measures will be reinforced.

For this hypothetical example, psychological reinforcement and procurement efficiency can be both easily, and yet importantly, served. Bombs that are twice as effective should permit fewer airplane sorties to be flown to accomplish a given task, and at the least the likely savings in very expensive planes must be estimated as well as the relatively trivial addition in bomb procurement costs. A first approximate measure of this savings is likely to reverse the cost/effectiveness finding decisively, and lead to a decision that the new bomb will replace the old. Sound military intuition will be confirmed. Or reverse the standard test, holding a total budget constant while permitting aircraft funds to be diverted to the improved bombs as long as this substitution enhances military capability, and the same result will obtain in a procedure that the economist finds analytically equivalent, but which has greater military appeal.

Finding such examples in practice, where carrying substitution analysis one level higher yields decisive results, is fortunate. But the easy substitution analysis may not be decisive, for the easily calculable first-order savings in complementary inputs made possible by the new ordnance may not be enough to offset the additional ordnance costs. Such cases, in principle, then raise the dilemma between analyses that are tractably narrow but suspect, and those that would be definitively thorough in scope if only they were practicable. To test whether a sufficient case for the new ordnance can be made, it may be necessary to consider the full offsetting savings that it would make possible in supporting facilities as well as aircraft. The counsel of perfection is to adapt future operational concepts and tactics that exploit the potentialities of the new bomb to their fullest; to redesign bomb-carriers, bases or aircraft carriers, and all other support facilities accordingly; and then, in a necessarily very extended, complicated, and uncertain analysis, see whether offsetting savings in nonordnance costs would be sufficient. Nor, in principle, would even such a large scope be adequate. If the redesigned air systems are very different and much better, should the balance between the air and the partly complementary ground and naval arms be reexamined? For this balance, do income effects from weapon substitution run counter to, and more than offset, substitution effects? For the counsel of analytic perfection, there is no end.

The central conceptual problem, to borrow both the name and substance from its notable exposition by Hitch and McKean,[3] is to determine a level of suboptimization that is neither too small nor imprac-

[3] C. J. Hitch and R. N. McKean, *The Economics of Defense in the Nuclear Age,* Cambridge, Mass., 1960, esp. Chap. 7.

tically large for the problem at hand. The related central task is to invent, design, and redesign alternatives, or combinations of alternatives, that will permit analysis to be carried in a practical way to higher levels of suboptimization, and yet yield dominant or near dominant results. Because these themes have been well developed elsewhere,[4] they need not be elaborated here. Among the professionals, the economist can claim a comparative advantage in grasping the conceptual problem, and he should exploit it. He knows that one cannot be grandiose, and put numbers in all the variables of a social welfare function. Yet from his Sophomore days at college onward he also learns that particular allocations may be determined to fall inefficiently within the transformation frontier. Directions for reallocation may then be safely found within the restrictions of Paretian optimality; or, where they exceed these restrictions, their distributive or other implications can at least be made explicit. These are appreciable conceptual advantages. The economist can claim no comparable comparative advantage in inventing or redesigning concrete alternatives, other than the sensitivity he ought to have about the goals for system design, although this may be enough to overcome any comparative disadvantage. Since on balance the economist has much to contribute to military as well as to other applied areas, he can expect his concepts to spread within the military profession. But he also has much to learn, especially where overexposure to nonmilitary examples may have overly habituated him to marginalism.

III

For military as well as nonmilitary applications, efficiency tests for equal returns at the margin may be relevant, and may suffice except for determining a budgetary scale in value terms. Transformation functions may be shown to have simplifying convexity properties; returns to scale may be shown to be decreasing over relevant ranges. But the perplexities for choice that increasing returns pose are more likely to arise in military applications, plus the complications inherent in adversary relationships. At the very least, cases that fall within a two-by-two

[4] *Ibid.*, and A. W. Wohlstetter, "Analysis and Design of Conflict Systems," in E. S. Quade, editor, *Analysis for Military Decisions*, Chicago, 1964.

matrix that reflects these considerations need to be considered: decreasing returns to scale vs. increasing returns on the one hand, and direct military counters in kind to enemy actions vs. indirect counters. Of the resultant four cases, the decreasing returns and direct counter case is obviously the easiest to handle, especially in the static degenerate case where enemy military counters can be ignored.

All of us tend to use the simplest case for expository purposes, and to exploit it in practical application whenever possible. The favored example, derived from its central role in strategic force calculations, involves the allocation of strategic missiles to fixed targets.[5] Where the subset of such targets is taken that comprises enemy cities, the problem assumes an especially straightforward form. Because no society has yet shown any appreciable disposition toward dispersing urban areas as a means of reducing vulnerability to nuclear attack, a dynamic formulation to allow for interaction over time between missile allocations and this subset of enemy targets as to number and location is not needed. Nor is such a formulation as yet much needed for the "hardness" of cities, given the pace of civil defense programs, although this assumption must be more tentative. Finally, for the moment, calculations that reflect extensive antiballistic missile (ABM) defenses around cities are probably not needed. The complications that such defenses introduce can be considered later, as an especially pertinent example for the future.

Given such a target list, assigning a fixed number of missiles among the cities so as to maximize desired damage can be taken as the first and recurring allocation problem. As the operational planner moves beyond the obvious assignments of at least one missile to each of a few key cities, he will face diminishing returns for each additional missile that is assigned. There will always be a less rewarding extensive margin, since a smaller city can be added to the list. Also there will always be an intensive margin, for an additional missile assigned to a large city will both increase expected damage if all assigned missiles arrive and hedge against the possibility that previously assigned missiles will not be reliable. An optimal assignment will find marginal expected returns roughly equivalent at each assigned target, although average re-

[5] For example, Hitch, *Decision-Making for Defense;* and Statement of Secretary of Defense Robert S. McNamara before the Senate Subcommittee on Department of Defense Appropriations on the Fiscal Year 1967–71 Defense Program and 1967 Defense Budget, p. 47.

turns will be higher for the large cities, with a presumption that any added missiles would be divided among them at lessened marginal returns.

Let the defender add antiballistic missile (ABM) local defenses around some of his cities, however, and the offense planner's problem will be complicated. The defender will naturally tend to cluster such ABM installations around cities in proportion to the value he places upon them, seeking to make an attack on each defended city no more rewarding, as measured by the offense in terms of average damage per attacking missile, than one on an undefended, less valuable city. If the defender succeeds, he will lessen expected over-all damage for any given number of attacking missiles. He will, incidentally, also offset what would otherwise be diminishing returns to scale over a range for missiles assigned by the offense to defended cities. If the attacker now disposes of the same total number of missiles, will he cultivate the extensive margin further by attacking more undefended cities? Or will he concentrate more upon defended cities (the intensive margin) in order to override the defenses by force of numbers, and consequently attack fewer undefended cities? He may do the latter, which then yields the paradox that the way to defend small cities is to strengthen the defenses of the large ones. Counterintuitive as this possibility may at first seem, it can be put in the economist's terms as an indirect income effect for the attacker, since he is forced to a lower level of damage that, like the direct income affect of a lower budget for missiles, will induce greater concentration. This effect may overshadow the substitution effect in its tendency to replace defended with undefended targets.

The interactions among missile offenses and defenses that are different in kind extend so much further in their complications, however, that it is best to consider other basic situations first. Clearly the ABM planner is dealing with decisions now that will alter defenses years from now. Consequently he will be less interested in how the enemy might react in terms of assigning a given number of attacking missiles than in how he is likely to react in terms of expanding his offense force or, more generally, altering its characteristics. Can he react in such a way as to negate the efficacy of the defenses? Is he especially likely to do so because the costs to the attacker will be markedly lower than the costs to the defender; will the terms of trade, as it were, be favorable to the offense? These will tend to be the dominant questions, although the problem of missile assignment against any given defense remains a perennial related one.

IV

Less novel and complex interactions between opposed forces than is the case for ABM defenses can better illustrate the most fundamental reasons for increasing physical returns to scale in military applications. Such interactions arise when forces are designed to meet their counterparts directly—naval fleet vs. fleet, army vs. army, or, more concretely still, fighter plane vs. fighter plane, where the anticipated situation takes the form of a duel. Then two powerful reasons for increasing return apply: First, any firepower advantages in combat are likely to yield an edge that is more than linear in proportion; second, any perceived advantage in size or quality is likely to be multiplied as it redounds favorably upon the relative morale of opposed forces. Because this second advantage is especially difficult to estimate—"the moral is to the material as three is to one," is a slogan, not a finding—it tends to be ignored in quantitative analyses, which put it at zero; while it is sometimes elevated to overriding status in nonbudgetary qualitative appraisals, as if valued near infinity. The latter approach is clearly untenable, but the deficiency in the former also needs to be noted.

In contrast, firepower advantages lend themselves to rigorous quantitative formulation. One early formulation remains preeminent, for at one and the same time it addresses itself to the central combat relationship, expresses it in mathematically convenient form, and, once clearly expounded, appeals to intuitive common sense. Lanchester's Law—"The fighting strength of a force may be broadly defined as proportional to the square of its numerical strength multiplied by the fighting value of its individual units." [6]—has subsequently been extensively criticized and modified in various ways for applications to particular combat operations. Yet it has survived as the simplest analytic expression for a basic tendency that has been shown to apply when tested against past combat engagements.

Mathematically, Lanchester expressed expected casualties over time

[6] Frederick W. Lanchester, *Aircraft in Warfare,* London, 1916, p. 48.

for opposed "blue" and "red" forces in differential equations for each, with quality and quantity for these forces as functional variables. He put in model form a dominant characteristic that technology had already long since established for battle—that accurate firepower could be brought to bear from sizable distances, and therefore that the firepower of many individuals or units might be brought to bear upon one enemy individual or unit. Hence the intuitive appeal of something approximating his "*N*-square Law." Where aggregate firepower would be directly matched in battle, a numerical inferiority would penalize one both offensively and defensively. If outnumbered two to one by forces of equal quality, for example, you would expect to lose about twice as many units in a first volley. Consequently the beginning of a second volley would find you outnumbered by more than two to one, and so on through subsequent volleys that would produce a growing disproportion of forces.

In its various modified formulations, Lanchester's Law has important long-term implications for allocations among alternative forces, although its immediate applications are more directly to military operations. For short-term strategic and tactical choices, it relates to the famed Principles of War: to Concentration of Force (or Mass) obviously, more precisely if more rigidly than in the standard military literature, since it measures the premium to be placed upon outnumbering the enemy at points of combat. Less obviously it relates to Maneuver, and thus to an entire range of subtle skills associated with the art of war. For if in the short-run the size of the force and associated materiel must be taken as given, the brute force method for overpowering your enemy by outproducing him does not apply, while methods to outwit him do. If by your deception, his lethargy, or whatever, he concentrates in the wrong places, while you so maneuver as to achieve decisive numerical superiority in the right places, over-all success may be achieved despite aggregate parity or inferiority in size of opposed forces. In focusing upon firepower as the key quantifiable variable, Lanchester's Law by no means disputes the evident sense in propositions that firepower must be made available at relevant locations (Mobility) and that to know which locations are relevant (Intelligence) is indispensable.

Or the law can easily be put in the context of another Principle of War that reflects codified common sense, and is closer to traditional economics, that of Economy of Force. Put the problem for an entire military campaign in terms of minimizing your expected casualties for given territorial or other gains, and, following Lanchester, this criterion will ideally be met by a series of engagements in each of which crushing numerical superiority is attained. Then, benefiting from the "*N*-square Law," your casualties in each engagement will be very low relative to the opposition, and should accordingly be low for the entire campaign. This ideal is worth noting because the economist, once persuaded that increasing returns are in fact likely to apply, will find it in accord with his maximizing principles. The military professional may not expect this accord from the economist, since he tends to equate economizing with parsimony and therefore with a tendency to give a combat operation no more than it "needs." Against this tendency, paraphrasing from the account of an observer, one famed Supreme Commander in World War II gave the apt retort in rejecting calculations for the bare minimum in required support from an associated service: "Gentlemen, we are not staging an athletic contest whose aim is an even contest for spectators. This is war, and I require all the support that you can provide for as long as this battle may last." Cavalierly as this statement dismissed opportunity costs for the support in question, it at the same time redressed overconcern with them.

For longer-term allocations, opportunity costs may pose real dilemmas for choice. If forward, local, mobile, defenses are sought in many places, military objectives will be ambitious since each of these adjectives exacts its resource toll. In total, the objectives may imply budgets that are higher than the governments involved are willing to provide. Across-the-board budgetary restrictions are administratively the simplest means to reduce proposed military capabilities to relevant total budget levels. Yet given nonuniform military returns to scale in different places, uniform budgetary cuts will be nonoptimal; more strongly, given increasing returns to scale for such defenses in several places, uniform cuts would be decidedly disadvantageous relative to selective cuts that sacrificed such defenses in one or more places in order to exploit such returns well in other places.

The clearest examples fit past rather than current possibilities, mostly for one overriding reason. Now, as at no time in the past, one technically feasible global strategy can at least be considered that is exempt from such dilemmas. Protected strategic forces whose missile capabilities extend to at least half-global range, as is clearly possible when large missile payloads can be boosted into still more demanding orbits around the globe, could be procured in quantities sufficient to cover the civil societies of all prospective enemies; and, more critically, military objectives might then be taken to be fulfilled. The first element of such a capability is already foreshadowed: "Thus, . . . the strategic missile forces recommended for the FY [Fiscal Year] 1967–71 period would provide substantially more force than is required for an Assured Destruction [civil damage] capability against both the Soviet Union and Communist China simultaneously." [7] If American security policy were to change drastically to regard such a capability as sufficient in itself to cover all contingencies, military planning would be reduced to the missile allocation and assignment problems that were discussed earlier. Diminishing returns would be expected almost everywhere, and tests for marginally equivalent results would suffice to determine solutions to these problems. Nor would there be any serious problem about scale. Since current strategic missile coverage extends this far, and covers additional missions as well—and yet the entire American budget for Strategic Offensive Forces accounts for only about one-tenth the total defense budget of the United States [8]—relatively small total resources are needed by a nuclear superpower to be well along the curve of diminishing returns for these capabilities everywhere.

Because American policy-makers seek far less violent and more flexible military capabilities as well, however, the old dilemmas can arise, especially for what are now called General Purpose (tactical) Forces. They now arise less clearly than was formerly the case, notably for pre-World War I surface navies. There, minus the complications posed for land battle by terrain, firepower-squared advantages were so apparent that Lanchester was led to an extreme observation: "It is questionable whether under any circumstances it can be considered sound strategy to divide the main battle fleet on which the defense of a country depends." [9] What clearly applied to tactics also governed procurement, conspicuously in the Anglo-German naval competition. Great Britain considered herself compelled to maintain battle fleet superiority, and to announce so unmistakably that such was her determined intent, despite

[7] Statement by Secretary of Defense Robert S. McNamara, p. 49.
[8] Excluding special Vietnam costs. *Ibid.*, p. 213.
[9] Lanchester, *Aircraft in Warfare*, p. 38.

indeterminately high possible costs, that Germany hopefully would be deterred from pressing toward or beyond naval parity. One military requirement was made paramount.

If Great Britain had considered herself equally threatened in the Pacific, her central choice problem would have been harder. An Atlantic fleet that outnumbered its main rival decisively if concentrated, say in 8:5 proportions, would, if divided in 4:5 proportions in each of two widely separate areas, be critically inferior. Such a disposition of an existing force could well be the worst of possible choices, with military strength nowhere strong enough to supply much confidence for firm diplomacy. How can such a situation be avoided? The straightforward military requirements answer might be to double total strength, and thus to seek preponderance in each area. Qualms would arise on the part of the government, however, not merely about initially doubling costs, but about whether its clear intent would serve to spur rather than to deter arms races in one or both of the areas. Rather than move to a markedly higher scale, it might choose military superiority in the one area and accept inferiority in the other as the best among remaining alternatives.

V

With the disappearance of the battleship, and the extension of indirect naval combat over the full air, surface, and subsurface range, modern counterparts for physically increasing returns in direct combat are more likely to be found in land defense against overt aggression. Those key adjectives for such defense—forward, local, and mobile—may well, if taken as governing American military policy over much of the globe, raise similar dilemmas. Global extension, of course, implies a large scale of effort, and raises the old issue: How are the resources to be allocated geographically? If divided among many locales, small peacetime forward deployments can nonetheless, of course, be made compatible with effective local defense if prospective enemy capabilities are likewise small.

Or, even if opposed capabilities are larger, economic and yet effective land defense can be implemented, broadly speaking, by exploiting one or both of two traditional methods: First, defending commanders may be allowed wide flexibility as to place and time for battle. Then, for example, they may choose if necessary to retreat temporarily and trade space for time. In so doing, they can fight at locations where the terrain naturally favors the defense and has been made to favor defense still more by engineering efforts. Then much of an attacker's firepower advantages may be nullified by superior protection for the defender, while the attacker's logistics are relatively strained. Meanwhile, in the time gained by such tactics, the potential can be gained for a compensating offensive as reinforcements flow to the theater. The great programmed increase in American airlift and sealift capabilities makes such reinforcement now especially pertinent: "We are now proposing an expanded airlift program which will provide by FY 1973 an equivalent thirty-day lift capability from West Coast airfields to Southeast Asia more than ten times greater than that available in FY 1961." [10] Or, second, defending commanders may try to adapt to a political denial of flexibility, when local governments refuse to allow the ebb and flow of battle to move widely over their territory, and insist instead upon an inflexible forward strategy. In such a case, commanders, knowing the place of battle, however much they regret enemy knowledge of it, can try ahead of time to make the terrain as unfavorable to the attacker as possible. They can plan an area rather than mobile defense, and create a fortified zone to implement such a defense.

Historical examples abound. Perhaps the most prominent case of the first type of defense against superior numbers is supplied by the German Army in its retreat from Russia and Poland in World War II, when inopportune standfast edicts from Hitler did not mar its efficacy. A pertinent case of the second type of defense is perhaps best supplied today by South Korea, where large ground forces are prepared to fight in a partly fortified zone whose defenses have been strengthened over the years since the cease-fire. These defenses are more pertinent than the much more prominent but unjustly maligned Maginot Line because they, as the Maginot Line was never designed to,[11] stretch the full width of the possible land front. Moreover, the decisive 1940 attack by the Germans not only outflanked the Maginot Line to the north, but the Western allies moved forward to meet it in a series of engagements on terrain that neither side had prepared. Thus this spectacular campaign illus-

[10] Statement of Secretary of Defense Robert S. McNamara, p. 125.
[11] Vivian Rowe, *The Great Wall of France*, New York, 1959.

trated neither type of prepared defense, but rather the head-on-collision that, with roughly 50:50 odds for evenly matched forces, comes closest to the old naval analogy. Therefore, it comes closest to the simple test that, for unevenly matched forces, places such premiums upon the firepower advantages that can be brought to bear.

Such battle tests might apply in some future contingencies. The most obvious cases in point are for combat situations that might develop along lightly guarded portions of Chinese or Soviet borders, while the most publicized case is NATO's Central Front in Europe. There, geographically forward defense has been put as a political imperative, and mobile defense as a military imperative.[12] Yet, as one crude quantitative perspective, this front extends about four times as long as that in Korea, if all the bends along the Iron Curtain and the Austrian border are traced for a literal forward defense, while it is guarded by only about 35 per cent more ready divisions.[13] Qualitatively, the NATO front lacks the elements of a fortified zone that have been built in Korea, while "to fight an effective mobile defense requires greater tactical mobility, more armour, more conventional firepower and better logistical support than is found at present . . . [and] some restationing is needed." [14]

Putting the Central European forces perspective this way deliberately overstates the NATO problem in creating an effective local defense. The South Korean standard is excessive, in all probability, for this different area, where the Warsaw Bloc also has its great problems. Above all, NATO outmans the bloc in military forces now,[15] and should be able to convert these inputs to meaningful output advantage. But how? And how can traditionally conservative military staffs be persuaded, and through them their governments, that it has been done? Affirmative answers clearly should be possible. Yet if one combines military-preferred all-mobile units for flexibility, which tends to maximize unit costs, with politically-preferred inflexibility as to inconvenient place of battle, which

[12] General Lemnitzer, address to Western European Assembly, *NATO Letter*, July–August 1963, p. 20.

[13] The Institute for Strategic Studies, *The Military Balance, 1964–65*, London, 1965, estimates 27 divisions for Allied Forces Central Europe (p. 13), and 18 divisions for South Korea's First Army, to which 2 U.S. divisions must be added (p. 30).

[14] Lemnitzer, *Nato Letter*.

[15] The Institute for Strategic Studies, *The Military Balance, 1964–65*, p. 41, puts total active armed forces at 5,843,500 for NATO, and at 4,425,000 for the Warsaw Pact.

will tend to generate military demands for more units, the joint resource demands for an acceptable answer will be high.[16]

Even given these conditions, acceptable answers need not extend to the full requirements that would be traditionally calculated for militarily "worst case" contingencies; i.e., where the worst combination of enemy surprise and mass in attack is assumed. In the now fashionable phrase, it is realistic to plan forces to meet "crisis situations," in which battle is likely to develop out of a more-or-less symmetric test of nerves over time, but where neither side is at all likely to risk reflex nuclear over-retaliation by a "worst case" attack. The probable time in developing crisis situations permits acceptable economies through total force levels that incorporate a lessened safety margin against an enemy lead in mobilizing and in reinforcing forward elements, and through greater reliance upon such reserve elements as can be promptly incorporated within effective units. In addition, the great programmed increase in American air and sealift capabilities makes forces in the United States much more quickly available for a NATO front, or any other front, and these forces have been strengthened. For these reasons, superimposed upon fundamental Western resource dominance, adequate forward local defense should be feasible in Central Europe.

Over-all frontal adequacy, nonetheless, remains the relevant test. To orient planning toward possible crisis situations, even if one imprudently considers that no single situation is likely to grow to involve more than a few divisions for a short period, by no means implies that total forces can be held to anything like such small dimensions. To have only small forces to mobilize at one place would leave a superior enemy free to outmobilize NATO at that place, or to bring pressure to bear at undefended places elsewhere. Either would worsen rather than improve bargaining positions in a crisis situation, especially under conditions where enemy numerical advantages would have multiplied effectiveness. For the most publicized area for possible local defense, consequently, marginal analysis that aims at enhanced efficiency for a given total capability or budget can be highly useful, but doubly insufficient. Resultant force redesign may still fall short of adequacy; and, if it does, governments may well lack incentives to improve local defenses, even

[16] For this author's discussion of a militarily more efficient alternative, and its political problems, see "Rationalizing NATO Strategy," *World Politics*, October 1964.

when dominant solutions can be argued. However much the analyst may question a "requirements approach" by military planners in this instance, and rightly deflate some critical estimates, he must respect its fundamental rationale.

Elsewhere the most apparent examples where forward local defense poses, or rather would pose, great resource demands, are for exposed countries that prefer nonalignment and self-defense. Potential allies, especially the United States, might nonetheless be interested in helping to provide local defense in some contingencies. But for a subset of these contingencies, hastily improvised defenses can hardly be expected to be forward. For noncommitted nations, such is the risk they choose to run in preference to others, and presumably the matter can be left there. To be almost equally summary about the position of America's allies outside NATO, allied preparedness in and behind South Korea appears to be unique. For the most part, however, others face a distinctly lesser threat, so that speedy American reinforcements could tip the local balance of forces still more drastically, and often others are sheltered behind an ocean barrier that counts vitally. And where possibilities must nevertheless be faced that considerable land space might have to be traded for time, the space in question is often far less densely populated than western Europe. Accordingly, less constraining inflexibilities need be politically imposed upon military planning. For most, although certainly not all such areas, these circumstances bring requirements for peacetime deployments, and for critical base and other infrastructure elements that would permit emergency reinforcement, down to moderate levels. But they do not eliminate such requirements, and therefore do not eliminate the possibility that harsh choices must be made among areas to defend.

VI

Other principal sources for increasing returns in military applications arise also from geographical concentration, although not in the same traditional sense, and from production economies of scale. Sometimes these can combine, as for the problem of antiballistic missile (ABM) defense. To defend only one large city locally with such a defense, to take an extreme case, makes little sense. An enemy can then shift his attack to undefended cities, intensify his attack on the defended city in one way or another in an effort to penetrate the defenses, or ground-burst nuclear weapons upwind from the city in an effort to inflict fall-out casualties within it. If he is free to choose the most promising among these options, the prospective defense contribution of ABM installations that are so localized may be small. Yet the installations are bound to be large and extremely complex, so that they can distinguish probable missile warheads from decoys, and moreover it cannot be much less expensive to buy an interceptor missile that must operate with extreme speed and accuracy than it is to buy an offensive missile. Further, combine small deployment scale with large installation size and inherent complexity, and with high interceptor unit cost, and it is manifestly uneconomic for many producers to compete for the small market. Even one producer, or consortium, can hope to lower unit costs only if huge research and development costs are spread over a sizable market.

Such has been the general perspective that has counselled for either large-scale ABM operational deployments or none at all, and which to date in the United States has resulted in a clear choice for none. These same factors continue to operate, although, judging from Secretary McNamara's recent testimony, they operate with diminished force.[17] On the demand side, defenses prospectively would be useful against more than one threat as nuclear weapons spread. Specifically, defenses would be technically more effective, in all probability, against a future unsophisticated Chinese attack, if deemed necessary against this threat. On the supply side, the Secretary speaks of "exoatmospheric" interceptor missiles that, as part of ABM complexes, might extend defense ranges significantly farther. If ranges are so extended that "local" can become "area" defenses, with overlapping coverage, an attacker's opportunities will be significantly narrowed. Against complete area coverage of the nation, he could neither find undefended cities to hit directly, nor hit defended cities indirectly by fall-out from undefended impact points. Consequently again there would be increasing returns to scale in combat, in this case as gaps in an area defense were closed.

Given the general scope and character of a program for national defense against nuclear attack, a strategic analyst can examine many trade-offs among and between components for an ABM system, civil defense, and bomber defense. Marginal rates of transformation can be

[17] Statement by McNamara, pp. 55–58.

calculated in terms of likely reductions in damage against alternative possible attacks, and assessed against marginal costs. Yet again, rewarding and necessary as these analyses may be, they will be dominated by answers to the big questions: Is a defense needed? Against whom? Should an addition be made that would necessarily be large in scale, like an ABM system?

Answers will be assessed in terms of their likely impact upon the strategic programs of likely enemies, with all the uncertainty that this procedure implies. Unlike the comparative simplicity of the pre-World War I naval race used earlier as an example, enemy counters would more likely take a different form. Their answer to a new defense will probably be a more sophisticated offense, rather than something as simple and comparatively verifiable as a matching defense. Nor are the enemy's partly predictable reactions the only ones to take prominently into account. Would a new defense strengthen the credibility of American nuclear guarantees in allied eyes because it promised to reduce damage in the event of nuclear war in the United States? Or would it impair the credibility of American guarantees because it seemed to demonstrate excessive American fear of nuclear attack, and remind allies of their greater exposure? Whatever the answers to these difficult general questions, they are highly relevant to very aggregative decisions; and these aggregative decisions, in turn, will make some marginal analyses irrelevant and others highly useful.

For the production economies of scale that would still seem to apply for ABM defenses, and for many other military applications as well, little need be said to an informed audience. For many expensive military items the additional number which must be produced to constitute a revolutionary technological change is minute when compared with a civilian item like an automobile. It was early recognized that a "learning curve" tended to apply to airframe production for this reason, with unit costs declining in a fairly predictable pattern as production expanded.[18] Today, concentrated sources of supply tend still more to predominate. One spectacular example is the F-111 (TFX), which in different versions is designed to be the advanced tactical fighter bomber for both the United States Navy and Air Force, a strategic bomber for the Air Force, and probably for the services of allied nations as well.

[18] Harold Asher, *Cost-Quantity Relationships in the Airframe Industry,* The RAND Corporation, Report R-291, July 1, 1956.

Small production runs and great technical complexity combine to create concentration. Or one could take examples from space programs. How such economic concentration can be made compatible with a prudent spreading of great technological uncertainties among many research and development competitors is a highly pertinent problem, but one that fortunately falls outside of the scope of this paper.

Here it suffices to affirm that prominent production economies of scale do apply with special frequency in military applications, and that they can join with combat conditions to create increasing returns to scale that demand wide-ranging nonmarginal analyses of military alternatives. Many especially challenging problems are consequently posed for military systems analysis. The problems demand very broad analyses that can take scale and complementarity factors into account, and yet that must be reasonably solidly based upon many component studies. Marginal and supramarginal trade-offs ought somehow to be concurrently or at least consistently appraised, which is a very tall order. None of us as individuals, or as members of but one relevant profession, can be complacent about our abilities to undertake such analyses; but economic and other relevant bodies of theory can certainly combine with military insights to provide a basis for improved choices.