5. A METHODOLOGY FOR THE MODEL-BASED ANALYSIS

This chapter discusses the methodology that guides the model-based analysis portion of the dissertation. Following Quade (1989), it closely considers four basic elements of policy analysis: the objectives, the alternatives, the impacts, and the criteria. The last element of Quade’s framework—the model—is thoroughly discussed in Chapter 6. As described in the following sections, the methodology articulates a multi-attribute utility approach, borrowing on the work of Hillestad and Davis (1998) and Thie et al. (2001). The substance of this chapter also draws heavily from the findings and hypotheses highlighted in the preceding four.

5.1 OBJECTIVES: PURPOSE AND SCOPE OF THE ANALYSIS

The basic analytical goal is to help the Italian Army identify the transition path—and the associated sequence of decisions—that is most likely to lead to a balanced enlisted and NCO force, taking costs directly into account. As discussed below, the "balanced force" concept is related to whether personnel are appropriately distributed by grade/category and year of service. The simulation model is used to identify alternative strategies for attaining this fundamental objective, with the additional requirement that the force should by 2020 reach the steady-state mix of 30,000 VFBs, 45,000 VSPs, 10,000 Sergeants, and 15,000 Marshals.

Examining the degree to which different strategies bring about a balanced force over time is critical for several reasons. For instance, abnormally large concentrations of personnel in certain grades or in a small set of years of service make force management difficult—promotion bottlenecks would most likely ensue, negatively affecting career opportunities. In turn, this could have a direct impact on effort supply and on the quality of personnel that chooses the military profession, as
mentioned in Chapters 2 and 3. Imbalances could also have an impact on personnel productivity and on the skill mix: Chapter 4 suggests that unit productivity may be lower than potential productivity in cases where the experience mix is largely skewed toward junior or senior personnel. Past research also indicates that the returns to experience may be limited in those cases where soldiers conduct low-skilled activities. An excessively senior (or junior) force, therefore, could lead to significant experience and skill mismatches, which in turn would have a debilitating impact on overall effectiveness. Finally, to the extent that personnel costs vary according to years of service, large concentrations in career-track cohorts could also have a significant cost impact.

**Identifying subordinate--and more precise--objectives**

How can one assess the degree of force balance? Indeed, such a concept could by its very nature be measured in several ways. A number of subordinate objectives can add specificity and facilitate the process of evaluating whether particular policy options lead to different degrees of “balance”. Taking the findings from the previous chapters and the priorities of Army planners into account, four specific objectives stand out as particularly relevant:

- **YOS balance**: minimize the spikes in the YOS distribution relative to a target profile, and their concentration in particular age groups;
- **Promotion opportunities**: promote individuals at an acceptable rate;
- **Terminal grade share of category**: limit the number of individuals in the terminal grade for each category (e.g., VSP4’s share of the VSP stock);
- **Distance from the steady-state**: minimize the distance between the actual and steady-state force mixes during the transition.

---

93 Personnel humps and valleys would also prove disruptive to other personnel management processes, such as training and occupational assignment.

94 Other specific objectives could be identified and examined by this analysis. However, the ones presented here appear to be both relevant for and sufficient in capturing the essence of the broader objective.
The first objective is directly concerned with YOS imbalances--the goal is to avoid or reduce disruptive humps or valleys in the experience distribution of the force, especially when such imbalances are concentrated in particular age categories (that is, when they are not evenly distributed throughout the relevant YOS range). The second objective instead specifies that opportunities for promotions should conform to an ideal or target rate, reflecting the assumption that promotion opportunities should neither be too high nor too low. Related to this is the third objective, which underscores the potential effort supply, skill matching, and productivity problems that would arise from having an overly large number of individuals in the terminal grade of each category. The final objective instead highlights the fact that, at any time during the transition, being closer to the steady-state objective is "better" than being farther away, *ceteris paribus*.

In essence, all objectives articulate a set of goals for different aspects of force balance. The outcomes of alternative transition strategies are compared to such targets in order to assess their performance. The procedure for identifying alternative strategies is illustrated in the following section; the methodology for measuring these objectives and comparing outcomes is detailed in sections 5.3 to 5.5.

### 5.2 ALTERNATIVE STRATEGIES

Identifying alternatives that most closely attain the force balance objectives is one of the core tasks of the model-based analysis. These alternatives are essentially different sets of policy interventions to be implemented during the transition out to 2020 and beyond. These mixes can easily be altered in the simulation model, since the user can manipulate a set of levers that (1) have been included in similar efforts in the past, such as attrition and retirement timing and/or (2) are identified by the government and military personnel regulations as explicit policy variables, such as the number of individuals selected for the career force every year. Table 5.1 includes a list of policy levers that are particularly relevant for each personnel category.
Table 5.1 Policy levers

Junior Enlisted (VFAs/VFBs)
- Attrition
- % of individuals renewing
- % of individuals exiting the force at the end of their term (to police forces and civilian world)
- % of VFAs able to become VFBs
- Number of individuals recruited each year as VFBs/VFAs

Career Enlisted/Sergeants
- Attrition
- % of individuals selected for career force
- S1 slots: as a % of the VSP2, VSP3, VSP4 stock
- VSP2&3/VSP4 share of S1 slots
- Retirement age/probability

Marshals
- Attrition
- Retirement age/probability
- Number of individuals who enter the terminal M4 grade:
  - Through the competitive process before 8 years in the M4 grade
  - At the end of 8 years in the M3 grade
- Number of individuals selected from VSPs and Sergeants category
- VSP/Sergeants share of Marshals slots
- Number of lateral entrants

As explained in Chapter 7, the default set of values for these levers was specified either by official Army policy, or with the advice of several Army planners. Running the model with default values essentially produces a "base case" estimate of how the Army personnel system would behave over time in the absence of additional interventions. The base case scenario (also illustrated in Chapter 7) serves as a useful reference point when considering alternative strategies (Walker, 1978).

Constructing and screening the alternatives

The process of creating and screening alternatives is especially critical when using a simulation that generates a large set of runs (Quade, 1982, p. 129). To strike a balance between the infinite number of potential policy lever variations and the need to produce useful and timely results, strategy creation and screening were undertaken using a two-phase "search" process.
Stage 1: First-level screening of policy levers

In the first phase of the search process, the leverage of each policy variable is tested by parametrically changing its values in multiple runs, and by verifying the impact of such changes on model outcomes. Since it is unlikely that all variables will have high leverage over the behavior of the system, one should be able to identify which individual policies show the greatest potential for shaping outcomes and meeting the objectives.

Stage 2: Combining the levers to craft alternative strategies

The high-leverage policy levers identified in stage 1 are then combined to construct full-fledged strategies. The combinations simultaneously vary several parameters, as well as the timing of such variations during the simulation. Even with a smaller set of policy variables, it is still be impossible to undertake an fully exhaustive optimization, or a full exploration of all potential permutations. Having said that, by the end of the second stage a large number of alternatives will have been explored, and from these it will be possible to focus on a smaller number of interesting ones (Quade, 1982, p. 134). The shortlist will then be scrutinized more closely, especially in terms of the outcomes (or impacts) they bring about (see section 5.3).

The choice of policy levers and the determination of a model boundary

As one can infer from the policy levers listed in Table 5.1, most of the analysis revolves around changing the policymaker's decision rules. Individual responses to these changes, measured by variation in behaviors such as enlistment, re-enlistment, and effort supply, are not

---

Quade writes that "Once can always think of trivial ways of modifying an existing alternative to obtain an entirely new variant. But it is also obvious that at some stage it may clearly not be worthwhile to do so; although differences may exist, they may be so insignificant that, within the precision of our model or of our data, their effect on the benefits or costs cannot be distinguished from previously considered alternatives. Moreover, many will be obviously dominated by others—that is, another alternative can be found that is better in some aspect (say, lower cost) but equally good in all other aspects" (Quade, 1982, p.129).
explicitly modeled. Recruitment and other processes are treated as exogenous—-that is, they can be changed parametrically, but they will not depend on individual responses to policy levers such as entry pay, advertising, and number of recruiters. This approach also implies that potential behavioral feedback effects (e.g., reduced enlistments resulting from reduced promotion opportunities into the career force) will not be included.

This decision can be justified on two main counts. Firstly, while looking at all aspects of a problem may be ideal, it is impossible in practice: "[t]o perform analysis, some considerations, usually many considerations, must always be left out" (Quade, 1982, p. 49). Only certain parts of a given problem can be explored in detail, and that while this approach is by definition incomplete, it can nonetheless identify strategies that constitute improvements over the current situation. Secondly, the decision to exclude behavioral effects is heavily affected by the uncertainty that surrounds such effects—especially their magnitude and their sensitivity. For behavioral impacts to be faithfully represented in the model, one would need to understand in much greater detail how individual behavior would be impacted by different policy changes. More information is needed on responses to changes in pay, promotion opportunities, and how such responses would vary according to personnel characteristics such as quality. These data can only be obtained using thorough econometric analyses or experiments, which constitute sizable research endeavors in and of themselves.

5.3 IMPACTS: ESTABLISHING THE MEASURES OF EFFECTIVENESS AND COST

Simply put, the impacts are the consequences of each strategy. The model gauges a strategy's impact by reporting outcomes for a set of measures of effectiveness (and cost) at each time period. The measures of effectiveness are directly linked to the four specific objectives mentioned in the first section of this chapter (see Figure 5.1).

96 Indeed, the specification of which variables to exclude or to assume as exogenous should take into account the data and knowledge that exists on the policy issue under review (Quade, 1982).
Figure 5.1 shows that the attainment of the specific objectives is assessed with more than one measure—for instance, promotion opportunities are reported for junior enlisted personnel seeking to enter the career enlisted force, and for career enlisted personnel seeking to enter the NCO force. This approach enables a more precise reading of performance of the system in different sectors; it also makes the assessment of performance easier, since the Army usually sets targets at the category and grade levels.

Each measures of effectiveness is described below; the last subsection includes a discussion of cost measures.

**YOS balance measures**

Two specific measures are used to capture YOS balance: (1) average squared deviations from a target YOS distribution, and (2) the share of "young" and "old" soldiers in the career force.

**Average squared deviations from a target YOS distribution**

YOS imbalances can be quantified by measuring the average squared deviations of an actual YOS distribution from a target distribution.\(^\text{97}\)

---

\(^\text{97}\) Section 5.4 describes how the ideal distribution is derived.
The process of calculating averaged squared deviations is relatively simple—for each relevant YOS, the square of difference between the actual and target distributions is computed; these deviations are then averaged over the applicable YOS range. Formally,

\[ \text{AvgSqDev} = \frac{\sum (X^a_i - X^t_i)^2}{n} \]  

where \( X^a_i \) is the actual number of individuals in YOS \( i \), \( X^t_i \) is the target number of individuals, and \( n \) is the range of YOS considered in the calculation. For instance, assume that an “ideal” distribution for 3 years of service is of 100 soldiers in YOS 1, 50 in YOS 2, and 25 in YOS 3. Assume further than in a particular year of the simulation, the actual profile is of 50 in YOS1, 100 in YOS 2, and 25 in YOS 3. Then the averaged squared deviation for that particular actual profile for that year is of 1,667, as show in Table 5.2.

<table>
<thead>
<tr>
<th>YOS</th>
<th>Ideal Profile</th>
<th>Actual Profile, Year t</th>
<th>Squared Deviation</th>
<th>Average Squared Deviations, Year t</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>50</td>
<td>2,500</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>100</td>
<td>2,500</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>25</td>
<td>25</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1,667</td>
</tr>
</tbody>
</table>

Averaged squared deviations provide a single-point estimate of the degree to which unusual humps and valleys affect the YOS distribution in a given year. Out of all the measures of effectiveness considered in this analysis, this represents the only case where deviations from an ideal or target case are squared—reflecting the assumption that larger imbalances are disproportionately more disruptive than smaller ones. Such assumption is in line with the opinion of most Army planners that were interviewed. It is also important to recognize that the number by itself has little meaning; however, it becomes useful when these values are compared across time periods and across alternative strategies (see section 5.4).
Career force distribution by age group

Average squared deviations gauge “spikeness,” but do not provide much information of how concentrations of personnel are distributed across the force. It is conceivable that two YOS distributions could have very similar averaged squared deviation scores but very different personnel management implications. To take such a possibility into account, this set of measures calculates the share of two arbitrarily-defined age groups (young and old) as a percentage of the total. Large deviations from the target share would be symptomatic of either peaks or valleys in specific portions of the YOS distribution.

Promotion opportunities measures

Promotion opportunities are assessed for those seeking to transition from the VFB to the VSP category, from the VSP to the NCO categories, and from the M3 grade to the Marshals' terminal grade, M4. For each, this is computed as the number of promotions divided by all those eligible for promotion. Other promotions that take place in the system are not considered since they rely on automatic time-in-grade rules.

Career force terminal grade share measures

The share of career force terminal grades is simply obtained by dividing the size of each terminal grade (e.g., VSP4) by the size of the corresponding category (VSP). Other terminal grades whose share is gauged for every year of the simulation include S3, M3, and M4. The share of the M3 grade is included because access to the M4 grade as a M3 is regulated with a yearly contest (see Chapter 3).

Distance from the steady state criteria

This measure calculates the numerical distance between the steady-state force structure by category (VFB, VSP, Sergeants, Marshals) and the actual size of each category. The distance is measured for every time period of the simulation.

---

98 Sergeant promotions to Marshals are likely to be extremely low under most reasonable assumptions, and were therefore not included as a separate measure of effectiveness.
Cost measures

The model calculates the yearly cost (in constant 2001 Liras) of the active-duty force as a function of grade and years of service. Active-duty costs are derived from the Italian Army’s pay tables, and include basic pay as well as the additional allowances discussed in Chapter 3. The incremental costs of eventual early separation programs are also calculated.\(^9^9\) Such cost has to be approximated, since early separation policies have never been implemented by the Italian military. A suitable potential benchmark comes from the U.S., where during the drawdown of the 1990s personnel were offered early separation incentives. The more popular of these was the Special Separation Benefit (SSB), which consisted of a one-time lump sum equal to 15 percent of basic pay multiplied by years of service.\(^1^0^0\)

Given the considerable lack of uncertainty, the analysis assumes that the Army will manage eventual early separations by offering a SSB-like lump sum of either 15% or 30% of maximum pay times YOS. These are likely to be conservative cost assumptions, but such conservatism is warranted in light of the uncertainty surrounding this estimate.

5.4 GAUGING EFFECTIVENESS AND COST USING CRITERIA

A criterion is “a rule of standard by which to rank the alternatives in the extent to which they achieve one or more objectives. As such, it provides a way to relate objectives, alternatives, and impacts” \(\text{Quade, 1982, p. 47}\). Each measure of effectiveness discussed above is accompanied by a target quantity, and strategies are compared according to how far they deviate from such target. Some of these

---

\(^9^9\) Retirement costs are instead excluded from the calculation. The model does not also include other cost impacts such as training and equipment costs as functions of the experience mix. To be sure, the total cost impact of different strategies could vary in significant ways, but assessing total costs is a complicated enough endeavor to deserve a separate long-term study. Qualitative assessments of how such the total costs could be affected by different strategies total costs are made where appropriate.

\(^1^0^0\) The other early separation program, called the Voluntary Separation Incentive (VSI), offered personnel an annuity payable for 2 times their years of service equal to 2.5 percent times basic pay times YOS. Both the SSB and VSI required affiliation with a reserve component. See Asch and Warner (2001).
targets can be easily deduced from government plans (e.g., the steady-state category mix); others instead have to be obtained from expert opinion within the Army General Staff.\textsuperscript{101} Given that the model will produce results for several time periods, criteria need to account for the fact that a strategy's performance on a given measure may differ significantly over time. What follows is a first description of the criteria that are employed for each measure of effectiveness and cost; these are further elaborated in Chapter 8.

**YOS balance criteria**

**Average squared deviations criteria**

The target YOS profile serves as the yardstick for the “spikeness” of the actual experience distribution. Such profile essentially assumes a uniform distribution of personnel for each category (or group of categories in the case of VSP and Sergeants). For example, if the VFB stock is composed of 30,000 soldiers and the length of the term is 5 years, then there should be approximately 6,000 VFBs in each YOS slot (1 through 5). Following a similar logic, Figure 5.2 below provides a rough approximation of the target YOS distribution for the enlisted (VFB and VSP) and Sergeant force, using the Army's size goals for each category in the steady-state.

\textsuperscript{101} The frame of reference for the evaluation of all measures of effectiveness is the government plan. That is, the analysis focuses primarily on how the Army could reach its steady state objectives in a way that maintains force balance--and not on whether the Army steady-state objective is "optimal." Evaluating different steady-states goes beyond the scope of this dissertation, since it would constitute a sizable effort in and of itself. That being said, the set of outcomes obtained from this analysis could certainly inform policy makers' judgments on the suitability of the current steady-state category mix (see Chapter 9).
It is important to note that these stylized profiles do not necessarily reflect the Army's expectations of what the experience distribution should be, especially if one takes into account the fact that (1) individuals can be promoted to other categories (e.g., Marshals), and (2) personnel at every YOS leaves as a result of attrition. At the same time, however, they complement other measures of force balance by signaling the presence of abnormal personnel concentrations. The persistence of such concentrations over time is gauged by recording performance at different time intervals.

Career force distribution by age group

For this measure of effectiveness, outcomes are judged according to the deviation between the actual share of “old” and “young” career personnel is compared to a target value (see Chapter 8 for a discussion of these and other target values). Deviations from the target are

\[\text{102} \quad \text{Nor do they necessarily constitute an ideal profile in terms of force readiness and productivity. As mentioned before, it is impossible at this stage to precisely ascertain what would be the “optimal” experience and grade mix.}\]

\[\text{103} \quad \text{Estimates of average squared deviations will also take into account the fact that the VSP and Sergeants categories will not have a full YOS distribution until the very last years of the simulation, since they begin the transition exclusively with young personnel. For these categories, those deviations resulting from the lack of personnel beyond the age of the oldest cohort will not be included. See Chapter 8 for more details.}\]
assessed at different times by averaging them over time periods (e.g. 2001 to 2010, 2011 to 2020, and 2021 to 2030).

**Promotion and terminal grade criteria**

Strategies are also evaluated according to their deviation—either negative or positive—from a set of promotion targets derived with the help of expert opinion. For instance, the Italian Army wishes to promote a third of each eligible VFB cohort into the career force every year; values significantly lower than 33% signal an excessive tightening of standards, while higher rates indicate a loose promotion policy. Similarly to prior criteria, performance is assessed over different intervals.

**Distance from the steady-state criteria**

The numerical distance of each strategy from the steady-state personnel mix is easy to compute, since the Army has clearly specified what such mix should be from 2020 onwards. The strategies are then compared on their deviation from the steady-state over time.

**Cost criteria**

The model computes active-duty costs and early separation costs for each year. The simulation then generates a 30-year cost stream, which is evaluated in terms of its Net Present Value (NPV). Formally,

\[
\text{NPV}_{\text{costs}} = \sum_{t=1}^{30} \frac{\text{Costs}_t}{(1+r)^t}
\]

where \( r \) is the discount rate. The discount factor used is 3.5%, or roughly the net yield of an Italian 30-year Treasury bond. The debate on which discount rate should be used to evaluate government programs has been raging in the economics literature for decades. However, several sources indicate that analyses involving constant costs should use the real Treasury borrowing rate on marketable securities of comparable maturity to the period of analysis.\(^{104}\) The nominal yield on an 30-year Italian Treasury bond was 6.09% in May 2001 (Blooomberg.com,

2001). One obtains the 3.5% discount rate by subtracting from the nominal yield the current inflation rate which was estimated at approximately 2.5% (Economist Intelligence Unit, 2000). To measure the sensitivity of the results to these assumptions, alternative NPV calculations were made using a range of discount rates.

5.5 COMPARING THE ALTERNATIVES: A SCORECARD APPROACH

The process of comparing strategies is seemingly problematic, since there is no way simple way of ranking the alternatives. Indeed, there are multiple measures of effectiveness over several time periods, often measuring outcomes with different scales. The analysis is also complicated by the fact that policy makers are likely to differ on the amount of emphasis placed on each measure--some may be more concerned with short-term impacts on effectiveness, while others may instead worry about the long-term costs of each strategy.

These potential problems can be addressed by adopting a multi-attribute utility approach, whereby effectiveness is disaggregated into several factors which are then weighted according to a given set of preferences (Keeney and Raiffa, 1993). The Dynarank decision support tool provides an intuitive way to present the results of policy analyses based on a multi-attribute utility approach (Hillestad and Davis, 1998). Dynarank displays the alternative strategies in rows, and the criteria against which the options are to be measured in columns. The intersection of the rows and columns contains the evaluation or score of each option against each measure. In turn, these scores are weighted to reflect policy makers' preferences. The weighted average of all scores produces a measure of aggregate effectiveness.

As discussed in the previous sections, total effectiveness is a function of subordinate and more precise measures. The scorecard reflects this hierarchical approach, and employs three distinct "layers" of effectiveness. The top level consists of the four specific objectives discussed in section 5.1: YOS balance, promotion opportunities, terminal grade share, and distance from steady-state. The mid-level is composed of the measures of effectiveness corresponding to each specific objective (as discussed in section 5.3), while the base-level represents
the performance of the measure of effectiveness over different time intervals. Figure 5.3 shows the structure of the scorecard—for ease of exposition, only the first top-level measure and those mid-level measures pertaining to enlisted personnel are included (see Chapter 9 for more details and actual scorecard output).

Figure 5.3 Sample scorecard

Each criteria contains a weight (which is initialized to 1.0 across the board). The aggregate-value column represents the product of the normalized weights on each measure multiplied by the corresponding effectiveness and summed. For every higher-level measure, therefore, aggregate effectiveness can be expressed as follows:

\[
AE(x_1, x_2, ..., x_n) = \sum_{i=1}^{n} w_i(x_i)
\]

where

\[
\sum_{i=1}^{n} w_i = 1
\]

In equation 5.3, AE stands for aggregate effectiveness, \( x_i \) to \( x_n \) are the subordinate measures of effectiveness, and \( w \) is weight attached to each measure. At the most aggregate level, this expression represents an estimate of total effectiveness, or overall force "balance." Costs

---

105 Weights are normalized when used so that weights of 1.0, 2.0, and 1.0 across a set of three measures would actually be transformed to 0.25, 0.5, and 0.25 when used to multiply the option values. See Hillestad and Davis (1998).
are included on the right-hand side of Figure 5.2. With both measures of cost and effectiveness for each strategy, it is possible to rank alternatives by effectiveness, cost, and cost-effectiveness.\textsuperscript{106}

**Using the scorecards to rank options**

Aggregate effectiveness is a function of the "score" for each criteria and the weight placed on each. Given the multiple scales adopted for the different measures of effectiveness, the scores need to be standardized on a common scale—the one adopted for this analysis has a minimum of 1 (representing the poorest) and a maximum of a 100 (representing the ideal outcome). To convert raw model results to this scale, the following normalization function is used:

\[
100 \frac{|\text{Actual}\_\text{Outcome} - \text{Ideal}\_\text{Outcome}|}{\text{Ideal}\_\text{Outcome} - \text{Worst}\_\text{Outcome}} / 99
\] (5.4)

Equation 5.4 takes the absolute value of the difference between the actual result for a given strategy on a given measure and the target value for that measure. This is the divided by the absolute value of the difference between the target value and the value of the worst-performing strategy (which is itself divided by 99 to normalize the range); the entire expression is then subtracted from 100. The normalization function guarantees that for each measure the worst strategy will be assigned the value of 1; it does not guarantee, however, that the best will receive the perfect score (in fact, even the best strategy can be a poor performer on an absolute level).

This index produces a linear scale, as shown in Table 5.3.

\textsuperscript{106} Equation 5.3 is essentially an additive value function (Keeney and Raiffa: 1993). This assumes that objectives can be formulated in a way that guarantees their preferential independence. Such assumption is legitimate given the purpose of the analysis. Indeed, Hillestad and Davis argue that this approach is warranted since Dynarank "is for exploratory and screening analysis of options which ultimately may need to be examined in considerably more depth with respect to outcomes, costs, combinations with other options, etc." (Hillestad and David, 1998).
Table 5.3 Raw vs. scaled scores for a hypothetical target and set of strategies

<table>
<thead>
<tr>
<th>Strategies</th>
<th>Raw Score</th>
<th>Scaled Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>25</td>
<td>17.5</td>
</tr>
<tr>
<td>C</td>
<td>30</td>
<td>34</td>
</tr>
<tr>
<td>D</td>
<td>35</td>
<td>50.5</td>
</tr>
<tr>
<td>E</td>
<td>40</td>
<td>67</td>
</tr>
<tr>
<td>F</td>
<td>45</td>
<td>83.5</td>
</tr>
<tr>
<td>G</td>
<td>50</td>
<td>100</td>
</tr>
</tbody>
</table>

Other scaling approaches were considered, including the log-log format used in the United Nation’s Human Development Report (UNDP, 1999). This format would have boosted the scores of strategies in the middle of the range (shown in Figure 5.4) and by doing so, it would have lowered the relative performance of strategies at the extremes of the score distribution. However, the log-log method proves unfeasible when the target value is zero, as in the case of those measures gauging deviation from the ideal YOS profile and the planned steady-state. Moreover, in most cases where it was possible to observe the impact of the alternative scaling function, results did not substantially deviate from those obtained with a linear approach.

**Figure 5.4 The impact of using different scaling functions**

To be sure, the impact of each score and therefore aggregate performance depends on the weights placed on each criteria, and for each layer of effectiveness. It is plausible, for instance, that some
strategies ranking near the top under one set of assumption would fare poorly in another. Given the subjective and abstract nature of weights, there is no attempt to seek a definitively “right” set of weights, but rather to explore how different assumptions and weightings affect the relative ranking of options. This process, described fully in Chapter 8, leads to multiple ranked lists which are in turn examined for commonality and differences (Hillestad and Davis, 1998). In future research on this topic, the actual preferences of policy makers could be identified by applying questionnaire-based methodology similar to that described in Thie et al. (2001).