September 2008

The value of assessing weights in multi-criteria portfolio decision analysis

Jeffrey Keisler
University of Massachusetts Boston, jeff.keisler@umb.edu

The value of assessing weights in multi-criteria portfolio decision analysis

Jeffrey M. Keisler
Associate Professor
College of Management
University of Massachusetts Boston
5th Floor McCormack Hall, Room 249
100 Morrissey Boulevard
Boston, MA 02125
jeff.keisler@umb.edu
phone: 617-287-7738
fax: 617-287-7877

The value of assessing weights in multi-criteria portfolio decision analysis

Abstract: *Analytic efforts in support of portfolio decisions can be applied with varying levels of intensity. To gain insight about how to match the effort to the situation, we simulate a portfolio of potential projects and compare portfolio performance under a range of analytic strategies. Each project is scored with respect to several attributes in a linear additive value model. Projects are ranked in order of value per unit cost and funded until the budget is exhausted. Assuming these weights and scores are correct, and the funding decisions made this way are optimal, this process is a gold standard against which to compare other decision processes. In particular, a baseline process would fund projects essentially at random, and we may estimate the value added by various decision processes above this worst case as a percentage of the increase arising from the optimal process. We consider several stylized decision rules and combinations of them: using equal weights, picking one attribute at random, assessing weights from a single randomly selected stakeholder. Simulation results are then used to identify which conditions tend to make which types of analytic strategies valuable, and to identify useful hybrid strategies.*

Keywords: *Multi-criteria decision analysis, portfolio.*

1. Introduction

Background

In portfolio decision analysis, decision analytic techniques are used to estimate the value of the different projects in a portfolio after which a set of proposed projects is selected to receive funding in order to maximize expected utility. This is a pervasive application. Because efforts to assess decision model parameters are themselves costly and require scarce managerial attention, it may be necessary to focus these efforts where they will be most effective and add the most value. Otherwise, as often happens in practice, the decision maker and analyst may rush through a very important analysis, or at the other extreme, analyze far beyond the point of diminishing returns.

Researchers have modeled the impact of different analytic methods in multi-criteria decision making and portfolio project selection, in some cases finding that simple heuristics approach the results of sophisticated optimization techniques. But often, the practical question is not which exact algorithm to use – computing costs are nil. Rather, decision makers need to know which of many available assessment strategies and problem structuring strategies to pursue and when, i.e., where to focus managerial attention and analytic efforts, as well as how much difference it makes, and how to justify these choices to clients. In the spirit of Phillips' (1984) requisite decision modeling, we wish to know where more and less intensive approaches are sufficient. With simulation and sensitivity analysis, this paper illuminates such problems in the context of multi-criteria portfolio decisions.

One technique to estimate the value added by decision analytic efforts is to treat the outcome of a planned analysis as a random variable and to, essentially, calculate the expected value of information for that variable with regard to the decision or decisions it informs (e.g., Watson & Brown, 1978). In particular, we shall follow the approach used by Keefer & Pollock (1980) and more recently and explicitly Keisler (2004). For a given portfolio and a given pre-analysis structuring of the problem, we consider the resource allocations that would be made in the absence of analysis and with the benefit of analysis. We quantify the increase in the portfolio value from the former to the latter and call it the value of (or value added by) analysis. This is consistent with the way the benefits of decision analysis are often marketed, where the value trajectory for a set of randomly ordered projects is juxtaposed graphically with the efficient frontier.

Multi-attribute techniques (Keeney & Raiffa, 1976) and multi-criteria techniques can be used to estimate values of projects in portfolios, e.g., Stewart (1991), or Henig and Katz (1996). In the relatively simple case of linear multi-attribute value functions, which we explore in this paper, the analysis consists of assigning weights to different attributes and assigning scores on the attributes for each project. Assuming linear values simplifies many calculations under uncertainty (Kirkwood, 1992, Corner & Kirkwood, 1996). In practice, this could be a reasonable assumption when the scale of the investments is small compared to the total resources available to the ultimate stakeholders, for the same reason expected monetary value can often be used for business decisions instead of more complex expected utility calculations (e.g., Smith, 2004). The weights represent the combination of scaling factors for the units involved and, more significantly, the tradeoffs between the different attributes.

Once weights and scores are assigned, and when costs for all projects are projected, our model stipulates that an efficient frontier is identified and an optimal allocation of resources is selected by picking the point on that frontier that exhausts the available budget (or, if costs and benefits are expressed in comparable terms and in the uncommon case that the budget is unlimited, projects are funded until the point at which the marginal cost matches the marginal benefit).

It can happen that a change in weights leads to a drastic change in what gets funded but only a small change in total utility – or that this change in what gets funded makes a big difference to the utility received by some stakeholders but not by others. Peerenboom et al (1989), for example, illustrated "performance curves" showing how different programmatic objectives benefited as the budget was increased, and these curves can fluctuate as weights are changed. Thus, the way in which these weights are set has a complex impact on the ultimate value of the portfolio from any given viewpoint. Stummer and Vetschera (2003) explore this problem further, identifying value trajectories when weights assigned by different individuals (or aggregates of these weights) are used in either centralized or decentralized project selection.

In other contexts, weights do not have much influence on ultimate outcomes. Keefer and Pollock as well as Keefer and Kirkwood (1978) considered specification of weights along with other parameters of utility functions for a handful of small portfolios, finding that it is most critical to incorporate non-linear utilities where appropriate and relatively less important in those cases to weight the attributes. For situations in which attribute scores are positively correlated (i.e., basketball player strengths), Einhorn (1976) found that the exact choice of attribute weights used had almost no effect on rankings. In a practical setting of IT budgeting, Jessop (2002) also found that project selection was non-volatile with regard to weights. Jia et al (1998) performed a detailed comparison of several weighting schemes for the problem of selecting one alternative out of several, finding that equal weights do not always perform well, but rank-ordered centroid weights based on only an ordering of attributes lead to much the same choices as do actual weights. Extending this work, Butler and Jia (1997) found a given real decision problem was not sensitive to the weighting scheme used. Roberts and Goodwin (2002) refined these approaches and found that ranking of attributes can be powerful. Liesio et al (2007, 2008) considered a richer portfolio including possible synergies between projects, and obtained bounds on the value of optimal solutions under incomplete information about weights. Thus, there is evidence that precise weights are not always needed. All of these studies involve a variety of strategies for assigning weights, and in some cases a variety of measures for the resulting performance of the portfolio.

To synthesize strengths of previous studies, we simulate a large number of decisions to achieve statistical results more robust than indications from single cases. We consider portfolio decisions specifically, rather than selection of a single alternative (where the impacts of characteristics such as variation among values, size of choice set, etc., are mathematically much different), while still comparing strategies using simulation somewhat along the lines of Jia et al (1998). Portfolio decisions typically involve multiple stakeholders, so we consider how strategies involving multiple individuals perform from various perspectives. Finally, we compare a range of strategies that an analyst could undertake in settings with a typical range of characteristics, and use sensitivity analysis to identify conditions favorable to the use of each strategy. This produces both heuristics for choosing methods as well as quantitative statements about the value added by different steps in the process. In particular, the value of information frame links this with Keisler's earlier work which considered the value of a portfolio under four generic analytic strategies: Random funding (which gave a lower bound), population level estimates of the distributions of project value, partial information about the value of individual projects, and finally complete information about individual project value (which gave an upper bound on the value added by analysis).

Our point of departure here is to consider more detailed intermediate strategies (rather than just at a superficial level of more or less information about generic project value), meaningful in the context of multi-criteria decisions, for obtaining and applying partial information about attributes and weights. Rather than just getting better or worse estimates of project value, we consider several avenues for richer scoring of projects, and consider the relative value added by these strategies in varying environments. In order to build a model that combines these various notions without becoming too complex, we shall assume a structure with varying parameters for all project and portfolio characteristics of interest and simplifying assumptions when possible elsewhere. Our aim is to learn about the extent to which, under varying circumstances, the process for accounting for multiple criteria and assigning weights must be inclusive, precise, and complete, but there are many decisions an analyst makes in the course of a project that are not addressed here.

Plan of study

First we consider some prior research for which project attribute scores and weights are available. Using summary statistics derived from data for these cases, we set representative – or at least reasonable – simulation parameters for distribution of attribute weights and estimates of those weights, as well as project scores along the various attributes. We then simulate a large number of portfolios using those parameters. We define several analytic strategies in terms of the decision rule as a function of both available information and the information to be acquired by doing analysis. The specific strategies range from a gold-standard for analysis, obtaining the best possible attribute weights and making funding decisions based on them, to a straw-man situation in which projects are funded at random, or, equivalently, by some method that does not correspond at all to the actual project values. In between are strategies where information of varying quality can be obtained about the attributes and their weights.

For each simulated portfolio, the decisions and corresponding ultimate outcomes and value received are calculated for each strategy. Random funding represents the case where there is no analysis, against which other strategies are compared. We vary input parameters over a range of values, and also consider targeted variations on the main strategies. The implications of these results *vis-a-vis* existing practice and understanding suggest recommendations for use of pre-analysis fact-finding and contingent approaches to analysis.

In the second section, we discuss previous studies on portfolio resource allocation involving simulated portfolios. We next describe a model of the portfolio process and the simulation procedure. We define several performance measures to be used (for single and multiple stakeholder situations), along with the main strategies to be considered. The next section presents simulation results for the various strategies for the representative base case, along with a basic sensitivity analysis followed by consideration of some new strategies inspired by the results. The last two sections describe a short application using these insights and then discuss the findings.

2. Model and assumptions

Structure

A portfolio consists of m candidate projects. Strategy k may recommend funding project i (F_{ik} = 1) or not funding project i (F_{ik} = 0). Projects are scored on a set of n attributes. For the purposes of this exercise, we assume there is one set of "true" or correct weights which would be revealed by a perfect process. These weights would yield value scores consistent with the decision maker's preferences over all alternatives involving tradeoffs across the different attributes. Weights can only be as true as preferences are consistent. In settings with multiple decision makers, preferences may not be consistent (Arrow, 1950) and so the construct of true weights is artificial. For other settings, the notion of "true weights" is intuitively clear. An extreme example is when every source of value can theoretically translate to monetary value, weights are just proxies for hard-to-calculate prices of the different attributes, e.g., the market share of a new product is valued by a firm exactly to the extent that it affects (in ways that are impractical to calculate) NPV of future cash flows. In still other settings, facilitating communication within a group allows them to agree to a single set of weights representing their combined interest (Phillips & Bana e Costa, 2007), which could be a simple aggregation of individual weights, or something more complex, and we can think of "true" weights as those the group would agree upon under the best possible conditions.

Each strategy produces estimated weights on all the attributes identified. One strategy might be to use a single individual's weights while a second strategy might be to use a different individual's weights. Project i's value given weights W_k identified under strategy k is assumed to be the product of the vector of weights and the projects' scores across attributes (j), i.e., for strategy k, $V_{ik} = \Sigma j\ (w_{jk}\ x_{ij})$.

In order to focus on complications related to the presence of multiple attributes, we assume in this paper that project attribute scores vary, but are already known and do not depend on the analytic strategy. (In the future, these results could be synthesized with similar results regarding other portfolio features.) We arbitrarily fix the cost of each project at 1, to avoid knapsack type complications where an unworthy project is funded merely to exhaust available budget at the margin. The results would not change significantly if costs are assumed to vary. Thus, strategy k maximizes ΣV_{ik} subject to the constraint that $\Sigma F_{ik} = C$, where C represents the budgeted cost available for the portfolio. Here, the number of projects to be funded is equal to C.

Scores for each project on each attribute are assumed to be drawn from the same distribution, where the natural log of the scores are normal with mean μ_x and standard deviation σ_x (which is consistent with proprietary empirical data obtained from two consulting groups). Attribute weights are formed in a hierarchical process (analogous to Lipscomb et al, 1998) where individual h's weight on attribute j is the product of the attribute's true weight w_{j0} and the individual's bias factor b_h, e.g., where individual h receives a noisy signal about w_{j0}, and where the log of the individual's perceived weight is equal to the log of the true weight plus a normally distributed error term. This assumption may be much too conservative for situations in which we actually expect negative correlations between different parties' weights (e.g., negotiations). These distributions characterize the portfolio at a general level, i.e., they don't tell us about weights of specific attributes, but by describing the level of uncertainty about those weights they are useful for information-acquisition decisions involving the assessment of those weights. More uncertainty would mean greater potential value added by the process that resolves it.

The w_{j0} are assumed to be independent and identically distributed (i.i.d.) variables where the log of the weights follow a normal distribution with mean μ_w and standard deviation σ_w (which implies attributes are defined so that weights are positive). This assumption is partly for computational convenience and partly based on empirical observation. The b_h are assumed to be i.i.d., where log b_h follows a normal distribution with mean μ_b and standard deviation σ_b.

Simulation parameter values

We choose reasonable attribute values without intending that they be interpreted as authoritative representations of the population of real portfolios – actual characteristics of real portfolios would obviously vary depending on their application areas, and hence it will be important to consider sensitivity analysis. Since the purpose of this analysis is to derive first-cut qualitative insights about the problem, rather than precise guidance on particular decisions, this should suffice. Simulations are based on the following parameters:

$\boldsymbol{\mu_x = 0}$

$\boldsymbol{\sigma_x = 1}$

This estimate for σ_x is based on analysis of proprietary data about project value (not attribute scores) from Strategic Decisions Group and Argonne National Laboratory, and the value for μ_x is arbitrary.

$\boldsymbol{\sigma_w = 1}$

This implies that the weight on an attribute in the 50th percentile of attribute weights is a factor of e (2.718) lower than an attribute in 84th percentile, etc. The value of 1 is a crude estimate based very loosely on reported data from existing studies. Kleinmuntz and Kleinmuntz (2001) had 8 attributes with weights ranging from 7.6% to 16.9%, suggesting a just slightly lower value of σ_w. Elsewhere, nine or ten point scales are common and this range of weights seems to have proven useful, suggesting a slightly larger value of σ_w.

$\boldsymbol{\mu_w = 0}$

The value of μ_w is arbitrary, since for our problem, it is simply a scaling constant and it won't affect the relative value of the portfolio under various different strategies.

$\boldsymbol{\sigma_b = 1}$

These parameter values imply that the variation in attribute weights due to differences among individuals (w_{11} vs. w_{12}) is of the same magnitude as the variation in weights across attributes w_{11} vs. w_{21}. Likewise, the variation between the contribution to project value resulting from achievements on different attributes is due in equal parts to variation in performance on each attribute and variation in the attribute weights.

$\boldsymbol{\mu_b = 0}$

i.e., we assume no systematic bias.

m = 500 (we need a fairly large portfolio to consider some of the problems. This is on the high end of what might be observed in practice).

C = 200 (a fairly typical portion – 40% – of projects are funded, which avoids problems of extremes where very few projects are funded leading to great variation, or where most of the projects are funded leading to almost no variation).

n = 10 (a round number just above Corner and Corner's (1995) average of 8.8 attributes).

We simulate attribute weights for two different individuals.

A Microsoft Excel spreadsheet simulates 500 portfolios using formulas based on unique random number seeds to generate attribute weights and the score for each project on each attribute.

Plan for sensitivity analysis

In addition to the base case, we shall consider results for various combinations of parameters, allowing n to be either 3 or 10, and σ_x, σ_w and σ_b to be either 0.5 or 1.0.

Measures and strategies

For each portfolio simulated, we consider the value $V_{(p)k}$ of a potential funded portfolio (p) according to the weights used with strategy k. Specifically, we consider the following valuation rules: when k = 0, $V_{(p)k}$ evaluates p using true weights; when k = 1 $V_{(p)k}$ evaluates p using (arbitrary) individual 1's weights; $V_{(p)2}$ uses individual 2's weights, $V_{(p)3}$ places 100% of weight on attribute 1 (attribute numbers assigned in no particular order), and $V_{(p)4}$ places 100% of weight on attribute 2; in addition to calculating the raw value, we also calculate this value as a percentage of the value of the 100% funded portfolio in each case (to illustrate the convexity of each value measure's performance curve). The comparisons between some of the individual values under different weighting schemes are directed at questions of organizational feasibility, e.g., predictability, stability and fairness.

We compute raw values and report normalized values, the former representing the case where the actual benefit-cost ratio for the projects may vary, the latter where the total of the weights sums to 1. We also calculate increments in raw values (above a baseline) arising from the strategies to control for variability in the simulated portfolios.

The strategies below do not represent an exhaustive list – in fact, a very large number of combinations of potential interactions and characteristics of projects could be included or excluded from a particular portfolio decision analysis, and real techniques can be used with varying levels of emphasis. The focus is less on detailed tactics than on drivers of portfolio value, along the lines of Guikema & Milke (2003). An alternative approach (e.g., Heidenberger & Stummer, 1999) is to model several specific algorithms exactly, in order to generate valuable insights about those algorithms – especially when one of those specific algorithms is under consideration. We define strategies that will be practical to simulate but still correspond to some of the choices analysts really face. The primary strategies are:

S0: Maximize value under correct weights ($\max_F V_{(p)0}$, s.t. $\Sigma F_i = C$);

This strategy represents an ideal decision analysis in which both quantitative assessments and group facilitation are successful. It serves as a benchmark against which other strategies are compared. We assume that there is a best set of weights that could be found for the decision and that this strategy determines those weights and then funds projects so as to maximize the value according to those weights.

S1: Maximize individual 1's value ($\max_F V_{(p)1}$, s.t. $\Sigma F_i = C$);

This strategy represents a decision analysis that does high quality quantitative assessments but ignores group facilitation. Instead, one person provides the definition of value used for funding decisions. In business, this could be, for example, this could arise if one manager's opinion of what's best for the company dominated the decision process. In social decisions, it could reflect a process that reflects one group's values while ignoring others' values.

S2: Maximize individual 2's value ($\max_F V_{(p)2}$, s.t. $\Sigma F_i = C$);

This strategy is the same as S1, only with a different (and *a priori* indistinguishable) person maximizing his or her definition of value. Defining S1 separate from S2 and $V_{(p)1}$ separate from $V_{(p)2}$ allows us to evaluate the consistency of certain results.

S3: Maximize score on attribute 1 ($\max_F \Sigma i\ (x_{i1})$, s.t. $\Sigma F_i = C$);

This strategy represents a process focused on a single issue that has no special claim to being the most important. It maximizes achievement on a single attribute while letting the others fall where they may. It is perhaps the simplest strategy that is better than no strategy at all, but it is not a strategy that an organization should consciously choose. Something like this could occur if individual psychology or group dynamics result in anchoring on a single issue which is then used as the basis for selection.

S4: Maximize score on attribute 2 ($\max_F \Sigma i\ (x_{i2})$, s.t. $\Sigma F_i = C$);

This strategy is equivalent to S3, and is included only to allow us to calculate various measures of consistency.

S5: All attributes weighted equally ($\max_F \Sigma_{ij}\ (x_{ij})$, s.t. $\Sigma F_i = C$);

This strategy has practical appeal. It is simple to apply, e.g., facilitation could consist mostly of brainstorming to identify all attributes. Scores are assigned and simply summed or averaged. No weighting step is necessary, but neither is any attribute neglected.

S6: Fund projects at random ($F_i = 0$ for $i > C$, 1 otherwise);

This strategy represents a worst case and is included to provide a baseline against which we can calculate how much better other strategies perform. Something like this can arise in practice when funding decisions are made on some basis uncorrelated with true value, for example, a first-come first-served funding process or one based on favoritism.

S7: Fund all projects: ($F_i = 1$ for all i)

This strategy is included in order to provide a denominator, in order to calculate the percent of total potential value achieved at different funding levels for different other strategies.

It is desirable for a decision process to lead to stable and predictable recommendations, in addition to just achieving high value. One measure of the stability of decisions is how much one individual's perceived value for the funded portfolio changes when some set of attribute weights is used other than those offered by the individual. Beyond stability in value, project managers also desire predictability as to whether their own projects are funded. Furthermore, the extent to which the desired funding from any particular viewpoint is not reflected in recommendations, we may anticipate greater implementation challenges. With this in mind, in addition to value added, we track the number of shared elements between the portfolios funded by each pair of strategies, i.e., $\Sigma_i F_i|S0 \times F_i|S1$, $\Sigma_i F_i|S0 \times F_i|S2$, …, $\Sigma_i F_i|S5 \times F_i|S6$.

3. Results

Base case

The base case results reveal several points. Figure 1 shows the incremental value added by each step of analysis as a percentage of the range from the worst case to the best case. In the simulation, maximizing on a single attribute (S3) alone adds on average 8%, to baseline value (this and other reported results are rounded to the nearest percent) but much less than the other strategies. Not only does S3 use the wrong weights. It completely ignores scores for other attributes. Using one individual's weights (S1) adds

40% to the portfolio value, which approaches the maximum value added (from S0) of 48%. However, if all weights are assumed equal (S5) for the prioritizing step, the portfolio is still worth 38% more than if funded at random (S6). These results are consistent with Edwards and von Winterfeldt's (1986) results on SMART techniques, in which inexact weights work rather well. The last result in particular is in line with Einhorn's results on equal weighting techniques. The equal weighting strategy clearly avoids the pitfall of focusing on any one attribute, and with far less organizational complexity (and assessment complexity), it performs only slightly worse than the strategy of using weights provided by any single individual. Therefore, the additional analytic effort of assessing one individual's weights may not be justified. The benefit of assessing all weights correctly is still significant – another 10%, and may still be worth doing depending on what system is already in place.

--- INSERT FIGURE 1 ABOUT HERE ---

Figure 2 indicates the efficiency of the portfolio under each analytic strategy, that is, the percent of the fully funded portfolio's value (S7) that each analytic strategy yields when 40% of the portfolio is funded. The performance curve for a single attribute is somewhat concave when its weight is included in the assessment (S0 and S1), yielding 50% of its maximum achievement when 40% of the projects are funded. A typical attribute performs better under the equal weighting strategy (55% achievement for 40%), but is linear, i.e., no better than the random strategy (as we would expect) when a single different attribute is the basis for prioritization. The attribute that happens to be the basis, however, is able to achieve almost 80% of its potential value from 40% of its potential funding. The differences between the average efficiencies of S0, S1, S3, S5, S6 and S7 are all highly significant (and, of course, S1 and S2 are essentially identical as are S3 and S4), as efficiency percentages have a standard deviation of about 3%, so 95% confidence levels on this statistic are within approximately +/- 0.3%. In essence, the better strategies are attempts to ensure that the primary attribute is the right primary attribute. This suggests a new, potentially strong and simple strategy: pick the most important attribute and prioritize on the basis of it.

Stability results: The portfolios are fairly stable under the first three strategies – 78% of the projects funded under S0 are funded under S1, which means an individual ought to expect some disappointments when projects are approved. The individual faces more disruption if another individual prevails: 70% of projects funded under S1 are funded under S2. Equal weighting (S5) is about as stable as S1 and S2, funding 75% of the projects funded under S0, and 69% of the projects funded under S1 and S2. Worse is when a single attribute strategy is used. In this case, stability is 48% with respect to S3 and S4, and the random funding strategy leads to the theoretical minimum stability of 40% (the proportion of projects funded).

--- INSERT FIGURE 2 ABOUT HERE ---

Sensitivity analysis

The results show that getting the attributes right is critical to quality resource allocation decisions, and getting the weights right is important but not as critical.

Reasonable shortcuts also add substantial value. If there is a different distribution over attribute weights, the results do not change radically, but they change enough that our preferred type of analysis could also change. We now consider results from four specific modifications on the base case:

1) Less variation among individuals' weights (σ_b = 0.5); In this case, because the individual judgments are representative, using a single individual's weights (S1) increases the portfolio value 95% as much as S0. Other strategies' effectiveness is unchanged; S1 remains significantly better than equal weighting (S5). 2) Less variation among weights (σ_w = 0.5); the overall improvement of S0 over random funding is 18%, less than half that in the base case. In this case, the true signal about weights is comparatively weaker compared to individual-level variation, so the value added by S1 declines to 75% of that from S0. Conversely, because weights are closer to equal than in the base case, the equal weighting strategy, S5, adds 90% as much value to the portfolio as S0. Use of a single attribute (S3) is better than in the base case, but not enough to make it attractive – still adding only 25% as much as S0. Distributions other than lognormal could be used, e.g., Keisler, 2004, compared results from the lognormal against those using a uniform distribution with the same standard deviation, but found only minor changes in the results.

3) Less variation on project-level attribute scores (σ_x = 0.5); The total increase in portfolio value due to S0 is only 22%, as opposed to more than twice that in the original simulation. The relative contributions of the other strategies, as a percentage of the value added by S0, are nearly unchanged.

4) Fewer attributes (n = 3); The total increase in portfolio value due to S0 is 38%, still large. We note two main differences here from the base case. First, all the strategies do well. Even S3 and S4 add between 75% and 80% as much value as does S0. S5 adds 96% as much as S0.

Extension to other possible strategies

The relatively strong performance of the equal-weighting strategy suggests that precise weighting may not be necessary, while the poor performance of the single attribute strategies suggests that we cannot avoid spending time evaluating project performance on multiple attributes. Can we still get closer to the optimal portfolio without conducting a full analysis? We now consider several new strategies that combine features of the ones just studied.

S8: Have the group identify the attribute with the highest weight and rank projects in order of their scores on that attribute.

This strategy – like S3 and S4 – requires minimal effort in scoring and only limited consideration of the relative weight of different attributes (i.e., it's necessary to know which one is most important, but other possible attributes do not even need to be identified). The group process could be as simple as a vote, or could be a more structured discussion. By focusing on the most important attribute, S8 assures that what is achieved will be useful. In some cases, rigorously identifying the most important attribute is not trivial – it may require specification of the relevant range of values for each attribute and then determining which attribute should have the highest weight given those ranges. In other cases, the ranges may be obvious and it should be straightforward to say that the highest weight should go on profitability or on lives saved, etc. Because the simulation model assumes all attribute scores are drawn from the same distribution, attribute weights proxy here for importance.

S9: Same as S8, but only use one individual's designated highest weight attribute.

This strategy is likely easier to implement than S8. The hope is that the individual will usually pick the group's most important attribute and even if not, it will still be an important attribute.

S10: Identify all relevant attributes. Identify the most important attribute and assign it a weight computed based on its rank. Assume that the normalized weight (out of 100%) of this attribute is at the statistical expected value of the highest weight attribute given the generating distribution for weights and the number of attributes. The remaining weight is split equally among the remaining attributes.

This strategy requires that all attributes be identified, but uses only minimal rank-based weights and does not require ranking other than picking a single most important attribute. It has the benefits of S8, and for the minor additional effort of identifying the other relevant attributes, it allows them to be considered as secondary factors. As with some of the other strategies, this strategy is conveniently compared with the others here, but in practice one might select more sophisticated approaches along the same lines, e.g., rank-ordered centroid weights.

Before comparing these strategies on full simulated portfolios, we first need to calculate the rank weights. In particular, we need the average weights for S10 and possible variants. In order to understand whatever difference there is between S8 and S9, it is also helpful to calculate how often these are the same and how much difference it makes. The table below shows averages calculated over 500 simulated weight vectors, for each of the sets of parameter values (number of attributes, σ_w and σ_b) listed. For each case, we calculate the average normalized weight of the highest weighted attribute, the average normalized weight of the highest weighted attribute as assigned by a single individual, the average normalized weight of the second-highest weighted attribute, and the proportion of the time that the individual's highest weighted attribute is the same as the true highest weighted attribute and the same as another individual's highest weighted attribute.

--- INSERT TABLE 1 ABOUT HERE ---

Returning to the base case, we find that over an additional 25 simulated portfolios S8 adds 65% as much value to the portfolio as S0 – much better than 16% for S3, and somewhat worse than 83% for S1 and 77% for S5; For S9 the figure is 44%, because it performs identically to S8 when the two strategies identify the same attribute as most important, and adds only 40% as much value as S8 (25% as much value as S0) when they do not match. We then compare S10 against the other strategies and find it performs very well: it adds 91% as much value as S0, it is much more consistently good than S1 and S5, and never much worse than either of them. Furthermore, S10 funds 83% of the projects that S0 does, so it could provide a palatable starting point for stakeholders whether or not further refinement of weights will occur. Figure 3 illustrates the value added to the baseline portfolio by (the information revealed by) these strategies.

It is possible to build on S10's promise with S10': After identifying the attributes, rank the top t attributes (instead of the top one) and do not actually assess their weights, but instead assume that their normalized weights are at their statistical expected values and split the remaining weight equally among the remaining attributes. As t approaches n, this strategy (or perhaps one using a centroid weighting scheme based on attribute ranks) resembles S0 – the attributes are identified and all estimated weights are much closer to their correct values than in an equal weighting strategy. Its performance will be very close to (but not quite equal) that of S0, but so will the amount of effort required to execute grows as well. A hybrid of these strategies is to apply S10 or S10' but to have a single individual (a benevolent boss) declare the most important attribute(s). With one top attribute, this strategy appears less attractive because the two different individuals identify a different top attribute most of the time. Having the individual rank several top attributes could lead to more robust results and quick assessment.

4. Discussion

In portfolios with multiple performance attributes it is crucial to make specific efforts to identify attributes rather than casually hoping that suitable attributes are used by default. It is also important to assess relative weights in some way, but not necessarily in exhaustive detail. The value of deriving weights is comparable to the value of information about project scores on the different attributes. The difficulty of the former depends on complexity of both the environment (number and complexity of attributes to weight and their required fidelity), and the organization (number of individuals to involve, the disparity of their viewpoints, and the degree to which they are ultimately agreeable), the difficulty of the latter depends mainly on the characteristics of the projects themselves (number of projects, complexity of scoring, required fidelity in scoring, and again number of attributes).

Use of equal weights may perform nearly as well as use of weights from a single individual, although if individuals tend to have similar weightings and attributes vary a lot in importance, a single individual’s weights will tend to better serve everyone than would equal weights. If a single attribute is to be used, it is critical that the attribute be the right one – the most important one. As long as the top one (or several) attributes are identified as such, there is limited downside to this approach. It is reasonable, intuitive and simple to focus on a single goal, though this is still not optimal. This is not robust if the wrong attributes could end up as the basis for choice; in that case the portfolio does little better than random. In addition to delivering high portfolio value, the equal weighting and perfect weighting strategies offer high stability in terms of which projects are funded, while other strategies lead to funding decisions that vary depending on which individual or which attribute is featured. Such stability, predictability, and perhaps fairness may provide an incentive for groups to agree to agree.

We have discovered one notable improvement on typical strategies: S10 requires decision makers to answer a single (albeit critical and non-trivial) question: which is the most important attribute? Then, using a reasonable off-the-shelf estimate for its weight and a simple equal weighting model for the rest of the attributes, we bridge half the gap between equal weighting and perfect weighting – with minimal effort. This SMART-style strategy is psychologically attractive, because any decision stakeholder ought to be able to state what they believe the most important attribute is, but as we move through the list of attributes it becomes harder to rank the remaining ones. In practice, similar heuristics exist. For example, some on-line decision support tools ask users to specify which feature of an object is most important and then provide rankings of objects.

Naturally, other factors also affect the attractiveness of strategies. There may be other organizational benefits to having stakeholders assess together a complete weighting. This could, for example facilitate implementation of downstream decisions or even creation of new project proposals (e.g., using Keeney's (1992) value-focused thinking approach). The current results should inform practice, but should complement and not replace the practitioner's real-world judgment.

5. Application

To illustrate how these findings may be useful, we review an effort in which they influenced development of a government R&D portfolio program. Here, funding priorities were to be set for about 100 projects falling into about 35 categories (between 1 and 6 projects per category) supporting 7 strategic objectives. Within each category, there were interactions among the projects e.g., one project's success – or failure – could reduce or eliminate the ultimate benefit from another project.

The initial plan was to develop an Analytic Hierarchy Process (AHP) type model by rating the importance of the goals, of the categories with respect to the goals, and of the projects with respect to the categories. A system to support this general type of analysis had already been approved and development started. Unfortunately, the developer had substantial expertise in system development but found some decision modeling issues to be intractable. While the top level objectives were well-documented, the AHP approach did a poor job discriminating among projects within a category level. The project level interactions were hard to characterize within an AHP structure, as were implications of probability of success in the face of these interactions. Also, the AHP approach's treatment of cost was not conducive to finding an efficient funding frontier. Realizing it needed outside help, the developer called in a subcontractor to aid in redesigning the architecture for the decision support system.

The main alternative under consideration was a multi-attribute utility (MAU) approach. But that had practical challenges as well. The number of projects was large, the number of attributes was large, and the number of people involved – even at the senior level – was large. An MAU approach would require sufficient training (of client personnel) to do sensible assessments. Conducting assessments with the right groups of widely dispersed people in the right order would have been quite time-consuming, especially because a lack of consensus on weights would hurt the credibility of the model. At this point, time and budget available for model development and analysis were quite limited.

By focusing on the value added by the portfolio analysis, it was possible to choose an approach that would provide the desired benefit to the client without missing development targets. Based on the insight that it is important to identify criteria and in particular to identify the most important criteria, and even better to give them somewhat higher weight, it seemed that the AHP approach using the already defined high level objectives and the relation of the categories to those objectives would suffice. Typically, one category of project was targeted toward one high-level objective. Therefore, more identification and more precise weighting of attributes would have been of limited value. At the lower level, however, project scores would have been unreliable because they failed to capture important interactions. Weights are only useful to the extent they are applied to accurate scores, so there was value to structuring a scoring system at the category level. The solution was to use a constructed scoring scale at the category level in terms of which projects were successful, and using the category weights already obtained, the system could then generate efficient frontiers. Success probabilities were then incorporated in a simple manner so that expected values could be used. Finding the funding allocation that maximizes expected value is then a fairly simple computational task. This hybrid approach promised a feasible cost of assessment and left a tractable modeling and computational problem. With robustness to weights and more accurate scores, this system would deliver nearly all the potential portfolio value.

6. Summary

The value of improved estimates on weights is analogous to the decision theoretic value of information about those weights, where possible experiments that yield information correspond to different choices about which assessments the decision analyst should conduct. This notion can guide the dialog between analyst and decision maker in determining the appropriate level of various efforts. The contribution of this paper relates the author's other work (e.g., Keisler, 2004). As in other models, here we have generic analytic strategies ranging from no information to historical background (knowing the attributes) to qualitative estimation (identification of important attributes) to quantitative estimation (specifying weights at lower or higher levels of precision). The contribution here is that we define the situation in more detail with respect to weighting in multi-attribute problems, with corresponding problem characteristics and potential analytic steps. The insights gained from this approach can be the basis of an explicit – though qualitative – pre-analysis step along the lines suggested by Butler and Jia, and as described in the application to design of a real portfolio decision process.

Under reasonable assumptions, our model suggests that the value added by improved weight assessments can range from nearly zero to over half the baseline portfolio value, comparable to reported value added from other types of analysis. Knowing the size of this range (and how sensitive it is to situational characteristics) is more helpful than merely knowing that more precise assessments are better. Each step in the assessment process can, under some circumstances, add substantial value. Thus, choosing which steps to include –and which to exclude – is not trivial. By considering first the characteristics of the portfolio (i.e., variation among projects), the organization (i.e., people and their level of differentiation), and the environment (i.e., the variation among attributes), it is possible to identify the analytic strategies likely to add the most value to the portfolio for the effort required. In practice, analysts should as a first step consider potential value added based on how the rough distribution of weights compares with those considered in the model, and then balance the results against the implicit difficulty (or cost) of different means of assigning attribute weights and prioritizing projects. Without pre-analysis, time-constrained analysts may opt for the efficient tactic of ranking one or more top attributes and estimating attribute weights based just on this information as described at the end of section 3.

The current model has limitations that future research should address. A richer model for valuing potential analysis would consider the impact on strategies when functions aren't linear, when projects or attributes are not independent, and when the organization has more complex structure than just one or two individuals with unpredictable differences of opinion. In addition, the methods available from which the multi-criteria decision analyst chooses are far more subtle and complex than those modeled here. Refinements of the model could certainly address distinctions between methods for estimating weights under different assumptions about how characteristics would present. Furthermore, as multi-criteria methods can be very dynamic, there are analytic choices that depend on the value of additional iteration and interaction among participants. Finally, the model could consider the efficacy of such methods in the face of strategic behavior in assigning weights when there are conflicting objectives. Perhaps these steps could advance the main idea of this paper from a model that produces general insights to a formal pre-analysis tool with which to select an optimal method.

Until then, practitioners can use the simpler insights from the current model to apply their efforts effectively. They should do something about weighting, but maybe should not worry about it so much.

References

Arrow, K. 1950. A Difficulty in the Concept of Social Welfare. *Journal of Political Economy* **58**(4): 328-346.

Butler, J.C., Jia, J. 1997. Simulation Techniques for the Sensitivity Analysis of Multi-criteria Decision Models. *European Journal of Operational Research* **103**(3): 531-545.

Corner, J.L. Kirkwood, Craig W. 1996. The Magnitude of Errors in Proximal Multiattribute Decision Analysis with Probabilistically dependent attributes. *Management Science* **42**(7): 1033-1042.

Corner, J.L. and Corner, P.D. 1995. Characteristics of Decisions in Decision Analytic Practice. *Journal of the Operational Research Society* **46**(3): 304-314.

Einhorn, H.J. 1976. Equal Weighting in Multi-attribute Models: A Rationale, an Example and Some Extensions. *Proceedings of the Conference on Topical Research in Accounting.* New York University Press: New York, pp. 1-13.

Guikema, S.D., Milke, M. W. 2003. Sensitivity Analysis for Multi-attribute Project Selection Problems. Civil Engineering and Environmental Systems **20**(3): 143-162.

Heidenberger, K., Stummer, C. 1999. Research and Development Project Selection and Resource Allocation: A Review of Quantitative Modeling Approaches. *International Journal of Management Reviews* **1**(2): 197-224.

Henig, M.I. Katz, H. 1996.. R & D Project Selection: A Decision Process Approach. *Journal of Multi-Criteria Decision Analysis* **5**(3): 169-177.

Jessop, A. 2002. Prioritisation of an IT Budget within a Local Authority. *Journal of the Operational Research Society* **53**(1): 36-46.

Jia, J, Fischer, G.W., Dyer, J.S. 1998.. Attribute Weighting Methods and Decision Quality in the Presence of Response Error: A Simulation Study. *Journal of Behavioral Decision Making* **11**(2): 85-105.

Keefer, D.L., Kirkwood, C.W. 1978. A Multiobjective Decision Analysis: Budget Planning for Project Engineering. *Journal of the Operational Research Society* **29**(5): 435-442.

Keefer, D.L., Pollock, S.M. 1980. Approximations and Sensitivity in Multiobjective Resource Allocation. *Operations Research* **28**(1): 114-128.

Keeney, R.L. and Raiffa, H. 1976. *Decisions with Multiple Objectives*. Wiley, NY.

Keeney, R.L. 1992. *Value-Focused Thinking* Harvard University Press: Cambridge, MA.

Keisler, J.M. 2004. Value of Information in Portfolio Decision Analysis. *Decision Analysis* **1**(3): 177-189.

Kirkwood, Craig W. 1992. Estimating the Impact of Uncertainty on a Deterministic Multiattribute Evaluation. *Management Science* **38**(6): 819-826.

Kleinmuntz, D.N., Kleinmuntz, C.E. 2001. Multiobjective Capital Budgeting in Not-For-Profit Hospitals and Healthcare Systems. Working Paper.

Liesio, J., Mild, P, Salo, A. 2007. Preference Programming for Robust Portfolio Modeling and Project Selection. *European Journal of Operational Research* **181**(3): 1488-1505.

Liesio, J., Mild, P, Salo, A. 2008. Robust Portfolio Modeling with Incomplete Cost Information and Project Interdependencies. *European Journal of Operational Research* **190**(3): 679-695.

Lipscomb, J, Parmigiani, G, Hasselblad, V. 1998. Combining Expert Judgment by Hierarchical Modeling: An Application to Physician Staffing. *Management Science*, **44**(2): 149-161.

Peerenboom, J.P., Buehring, W.A., Joseph, T.W. 1989. Selecting a Portfolio of Environmental Programs for a Synthetic Fuels Facility. *Operations Research*, **37**(5)**:** 689-699.

Phillips, L.D. 1984. A Theory of Requisite Decision-models. *Acta Psychologica* **56**(1-3): 29-48.

Phillips, L.D., Bana e Costa, C.A. 2007. Transparent Prioritisation, Budgeting and Resource Allocation with Multi-Criteria Decision Analysis and Decision Conferencing. *Annals of Operations Research*. **154**(1): 51-68.

Roberts, R., Goodwin, P. 2002. Weight Approximations in Multi-attribute Decision Models. *Journal of Multi-Criteria Decision Analysis* **11**: 291-303.

Smith, J.E. 2004. Risk Sharing, Fiduciary Responsibility, and Corporate Risk Attitudes. *Decision Analysis*. **1**(2): 114-127.

Stewart, T.J. 1991. A Multi-criteria Decision Support System for R & D Project Selection," *Journal of the Operational Research Society* **42**(1): 17-26.

Stummer, C., Vetschera, R. 2003. Decentralized Planning for Multiobjective Resource Allocation and Project Selection. *Central European Journal of Operations Research* **11**(3):253-279.

Von Winterfeldt, D., Edwards, W. 1986. *Decision Analysis and Behavioral Research* Cambridge University Press: Cambridge, UK.

Watson, S., Brown, R.V. 1978.. The Valuation of Decision Analysis. *Journal of the Royal Statistical Society A*, **141**: 69-78.

Figure 1: Value added to portfolio by assessments

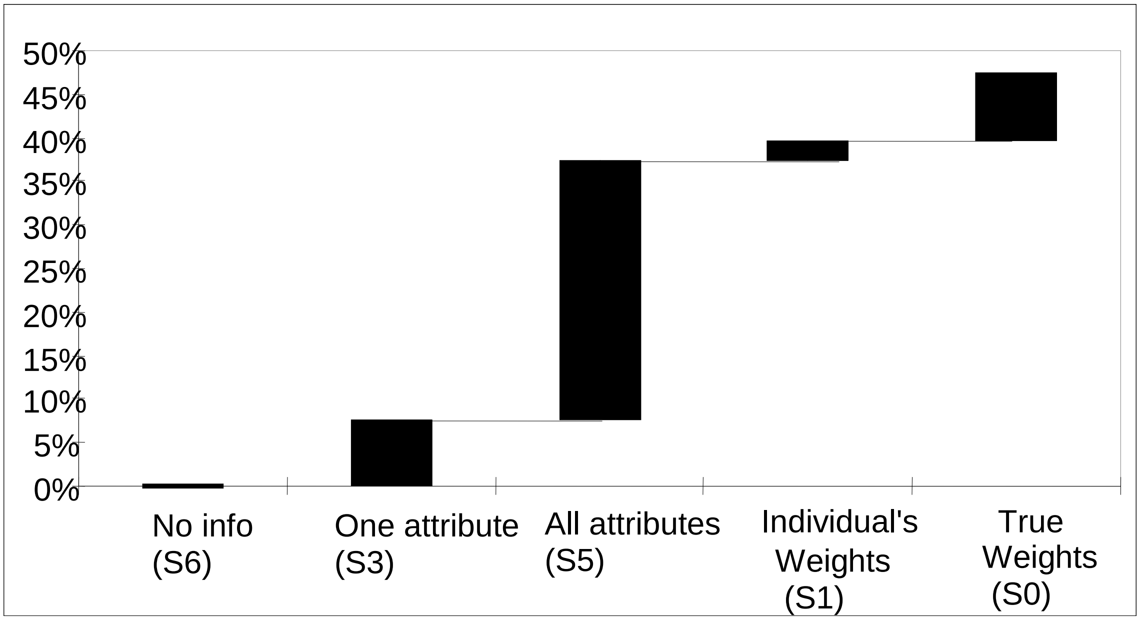

Figure 2: Percent of potential value achieved at 40% funding

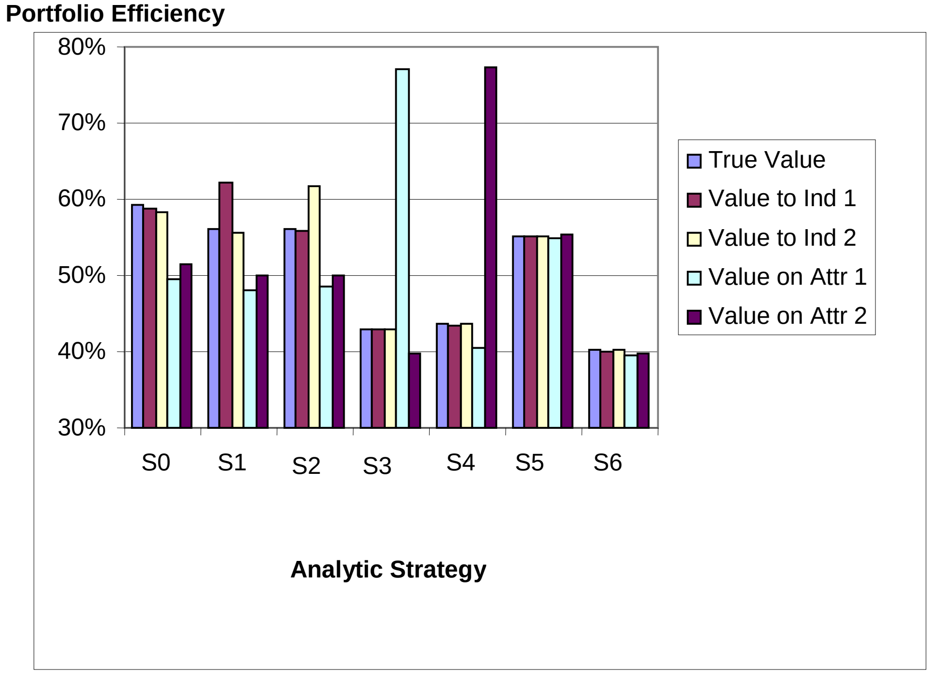

Figure 3: Value added to portfolio by frugal assessment strategies

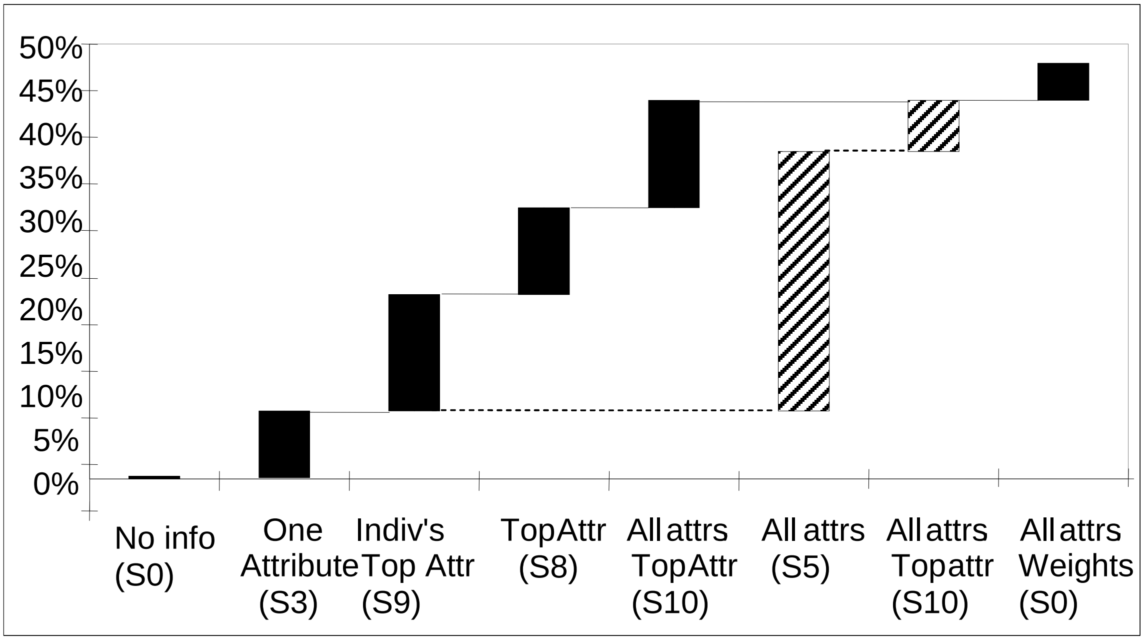

Table 1: Statistics on the highest weights

Attributes	σ_w	σ_b	Maximum weight	Maximum weight for indiv 1	Second highest weight	True vs. Indiv 1 matches on first attribute	Indiv 1 vs. Indiv 2 matches on first attribute
3	0.5	0.5	48%	53%	31%	*65%	53%
3	0.5	1.0	47%	62%	31%	51%	41%
3	1.0	0.5	61%	63%	27%	79%	71%
3	1.0	1.0	60%	67%	27%	65%	53%
10	0.5	0.5	20%	24%	15%	*39%	29%
10	0.5	1.0	19%	35%	15%	24%	16%
10	1.0	0.5	32%	35%	18%	57%	51%
10	1.0	1.0	32%	43%	18%	*39%	29%
* when $\sigma_w = \sigma_b$ the proportion of matches is theoretically equal for low and high values of σ							