Resource Allocation and the Value of Information

J. Eric Bickel
The University of Texas at Austin

Kun Zan
The University of Texas at Austin

Abstract: Value of information (VOI) is one of the most important and interesting applications of decision analysis. Yet, despite much research, few general properties of VOI have been proved. Intuitive properties such as the value of information increasing as uncertainty is increased have been shown not to hold in general. In this paper, we consider VOI within the context of a stochastic knapsack problem. We show that within this setting, some interesting and sometimes intuitive properties do hold.

1. Introduction: The Department of Homeland Security's budget in FY 2007 was approximately $43 billion [1], some of which supports the National Infrastructure Protection Plan (NIPP), whose goal is to "build a safer, more secure, and more resilient America by enhancing the protection of the Nation's critical infrastructure and key resources (CI/KR)" [2]. The NIPP decision-making process is based on the risk-management framework depicted in Figure 1. Its core is a risk-based resource-allocation phase (consisting of the steps Assess Risks and Prioritize) that is fundamentally a portfolio-management exercise or knapsack problem (KP)—even if not formalized as such. However, instead of the classic, deterministic knapsack problem, the policy makers in this situation face tremendous uncertainty regarding the probability of a significant event (e.g., terrorist attack, natural disaster, or other incident), the consequences of such an event, and the effectiveness of counter measures.

Figure 1: National Infrastructure Protection Plan's Risk Management Framework

These questions can be analyzed through the application of value-of-information (VOI) techniques, which seek to understand how much better off a decision maker is by gathering additional information instead of acting on his/her current understanding. VOI techniques can play a valuable role in the policy process, as the Presidential/Congressional Commission on Risk Assessment and Risk Management [3] concluded:

"In those cases where the quality of information is poor and the stakes in decision making are large, agencies should experiment with formal value-of-information methods to determine whether it is most appropriate to act or wait for improved information. Continued research in the methodologic development and application of value-of-information techniques…should be encouraged."

Despite the ubiquity of knapsack and VOI problems, neither type is routinely treated as such. There are several reasons for this. First, most real-world decision problems include a high degree of uncertainty and project dependence, which are difficult to model efficiently. Second, VOI analyses require detailed probabilistic assessment and modeling, which can result in complex Bayesian calculations.

1.1 Research Objective: To address the questions posed above and the lack of integration between knapsack and VOI problems, we are researching the properties of information value in the stochastic knapsack problem (SKP). Specifically, we have the following research objectives.

Objective 1: Integrate VOI and knapsack methodologies to better understand the drivers of information value in resource-allocation settings. This includes the impact of changes in knapsack capacity, the number of items, uncertainty in item values, accuracy of the information-gathering program, and knapsack variety (e.g., 0-1, bounded, unbounded).

Objective 2: Prove general statements or theorems regarding information value in knapsack settings (e.g., VOI monotonically increases with the number of items under consideration).

Objective 3: Disseminate the tools and techniques required to effectively integrate VOI and knapsack problems, through journal publications, conferences, and education.

In this paper, we investigate VOI properties with respect to parameters of SKP such as knapsack capacity, the number of items, prior uncertainty, and information system quality.

2. Background:

2.1 The Knapsack Problem: We consider the well-known knapsack problem and several of its variants. We are given a set of N items, each with a known weight w_i and value v_i. Let the vector of weights be $\mathbf{w}' = (w_1,\ldots,w_N)$, where the prime denotes the transpose, and the vector of values be $\mathbf{v}' = (v_1,\ldots,v_N)$. We seek a vector of item counts $\mathbf{x}' = (x_1,\ldots,x_N)$ whose total value $V = \mathbf{v}'\mathbf{x}$ is maximal, but whose total weight $W = \mathbf{w}'\mathbf{x}$ does not exceed some capacity C. Formally, the KP is

$$\max_{\mathbf{x}} \ \mathbf{v}'\mathbf{x}$$
$$st \ \ \mathbf{w}'\mathbf{x} \le C \qquad (1)$$
$$x_i \in \Omega,$$

where Ω is an activity-set constraint that governs the number of each item the decision maker (DM) can include in the knapsack.

Several variants of the KP, which differ in their activity-set constraint, are widely studied. For example, if Ω is the set of non-negative integers, then the KP is referred to as the unbounded knapsack problem. If x_i is instead restricted to be in {0,1}, then the 0-1 knapsack problem is obtained. If x_i is restricted to be in {0,1,…, b_i}, where b_i is the bound on item i, then we have the bounded knapsack problem. If the item value is equal to its weight, that is, $v_i = w_i$, then we have the subset sum problem. The introduction of additional resource constraints or multiple knapsacks yields the multidimensional knapsack or the multiple knapsack problem, respectively.

In many applications of the KP, the item values are not deterministic, but are instead random variables $\tilde{v}_i$ with known probability distributions $f(\tilde{v}_i)$ and joint distribution $f(\mathbf{v}')$. In this case, the deterministic value vector $\mathbf{v}'$ is replaced by the random vector $\tilde{\mathbf{v}}'$. This version of the KP is referred to as the SKP. In this case, maximization of the knapsack's total value loses its meaning and the objective function in (1) is replaced by a preference-ordering criterion such as maximizing the probability that the knapsack's value exceeds a particular target [4-6] or by a criterion involving first-order stochastic dominance [7-9].

We assume the decision maker has a von Neumann-Morgenstern utility function [10] defined over the knapsack's total value. In this case, the objective function in (1) is replaced by

$$\max_{\mathbf{x}} E[u(V(\tilde{\mathbf{v}},\mathbf{x}))] .$$

This objective may be nonlinear, yielding a nonlinear knapsack problem (NLKP). If the decision maker's utility function is linear in the knapsack's total value (i.e., the DM is risk neutral), then the objective function in (1) can simply be replaced by $\bar{\mathbf{v}}'\mathbf{x}$, where $\bar{\mathbf{v}}' = (\bar{v}_1,\ldots,\bar{v}_n)$ and $\bar{v}_i$ is the mean value of item i. We do not consider non-von Neumann-Morgenstern utility functions or objectives such as coherent risk measures [11] because we find the expected utility axioms to be compelling from a normative perspective. Future research could relax this assumption.

2.2 Value of Information: Assume that the value v of an item is a function of random variable $\tilde{s}$, with prior probability density $f(s)$, and of the action x taken by a DM. Since $\tilde{s}$ is a random variable, so is the value of the item, which we denote $\tilde{v}$. In the absence of further information gathering, the DM chooses his/her optimal action by solving

$$\bar{u}^* \equiv \max_{x \in \Omega} E_s\left[u\big(v(x,s)\big)\right] = \max_{x \in \Omega} \int_s u\big(v(x,s)\big) f(s) ds ,$$

where E is the expectation operator and $\bar{u}^*$ is the maximum expected utility.

Assuming the DM's utility function has an inverse, his/her certain equivalent, CE, is the amount given with certainty that has the same expected utility as the optimal action, which can be found by solving $\hat{v}^* = u^{-1}(\bar{u}^*)$.

Suppose that for an amount b, the DM is able to purchase an information system Θ that yields a signal θ regarding the outcome of $\tilde{s}$. The DM's expected utility for the system Θ is

$$\bar{u}_\Theta^*(b) \equiv E_\theta\left[\max_{x \in \Omega} \int_s u\big(v(x,s)-b\big) g(s\,|\,\theta) ds\right] .$$

The most the DM should be willing to pay for Θ, the value of information (VOI), is the value b' that solves

$$\bar{u}_\Theta^*(b') = \bar{u}^* . \qquad (2)$$

In the special case where the DM's utility function exhibits constant absolute risk aversion (CARA), that is, it is either linear or exponential,

$$VOI \equiv b' = u^{-1}(\bar{u}_\Theta^*(0)) - u^{-1}(\bar{u}^*) . \qquad (3)$$

Or, the value of the information system is equal to the certain equivalent of a costless information system less the certain equivalent without the system.

2.3 Literature Review: In this section, we review the most relevant literature regarding the value of information in single-project (or -item) settings in particular and in resource allocation problems in general.

Value of Information
The VOI literature has focused on developing general results regarding information value in the context of a single project, rather than in the portfolio or knapsack setting discussed here. In addition, most results merely prove that few general conclusions can be drawn. We highlight the main results below.

Flexibility: Hilton [12] proved there is no general monotonic relationship between the degree of action flexibility and information value. That is, adding (removing) actions to (from) the decision maker's feasible set does not necessarily increase (decrease) information value.

Risk Attitude: There is no general monotonic relationship between the degree of risk aversion and information value [12]. In the case of a binary accept/reject decision, Mehrez [13] proved that when the expected value of the accept alternative is less than or equal to zero, a risk-averse decision maker will never pay more for perfect information than will a risk-neutral decision maker. However, this relationship does not necessarily hold if the expected value of the accept alternative is positive.

Wealth: There is no general monotonic relationship between decision-maker wealth and information value [14-16]. In addition, LaValle showed that information value is independent of wealth if and only if the decision maker's utility function is linear or exponential. Where the decision maker's utility function is logarithmic, Morris [17] showed that VOI is a linearly increasing function of wealth.

Distributional Properties: Increasing uncertainty in the prior distribution does not necessarily lead to larger valuations of information [18]. Given a binary accept/reject decision and a risk-neutral decision maker, Mehrez and Stulman [19] demonstrated that the value of information is maximized when the expected value of the accept alternative is zero. Fatti [20] extended this result to the case of imperfect information.

Accuracy of Information: The VOI is non-decreasing in information-system accuracy, as codified by the likelihood function [21]. Clemen and Winkler [22] have studied the effect of dependence between information sources and found that even modest degrees of correlation between information sources can significantly lower the VOI.

Superadditivity/Subadditivity: The VOI is not additive across independent sources of uncertainty [23]. For example, the VOI on uncertainties X and Y together may be greater than (superadditive), less than (subadditive), or equal to the VOI on X plus the VOI on Y. However, for a two-act linear-loss decision with normal priors, Keisler [24] derived the conditions under which information value is superadditive.

Relationship between Perfect and Imperfect Information: In a single-project setting, Bickel [25] highlighted the importance of information accuracy by demonstrating that the value of imperfect information for an information system with correlation coefficient ρ is generally much less than ρ times the value of perfect information (VOPI). In addition, he derived the conditions under which the VOI is equal to $\rho \times 100\%$ or $\rho^2 \times 100\%$ of the VOPI.

Resource Allocation
The literature on resource allocation and capital budgeting is extensive, one strand being the knapsack problem itself [26; 27]. However, the literature regarding information value in the context of resource allocation or knapsack problems is relatively sparse.

Mehrez and Stulman [28] analyzed examples where a company can gain perfect information on a subset of projects at a specified cost. If the company faces an information-gathering budget constraint rather than a project-funding capital constraint, it was found that the company should obtain perfect information on the projects where it has the highest value.

Mehrez and Sethi [29] developed a hierarchical strategic-planning approach that integrates information gathering, project evaluation, and project funding. They addressed a company's facing constraints on how much it can spend on information gathering and how much it can invest in projects, but they did not analyze the characteristics of information value in this setting.

Keisler [30] analyzed the value of improved value estimates in the context of a constrained resource-allocation problem. He investigated several funding strategies in an attempt to isolate the sources of value in portfolio management. In particular, he sought to understand the value created by improved estimates of project value versus disciplined project-ranking criteria. Although Keisler's formulation was a 0-1 KP, with all projects having the same cost, he did not analyze the properties of information value in the knapsack setting.

Bickel analyzed the value of seismic information in oil and gas drilling decisions, modeled as knapsack problems [31]. He noted that information value seems to be maximized when the decision maker faces a budget constraint that is binding but not "too" tight. However,

the behavior of information value as a function of knapsack parameters was not analyzed.

In sum, the VOI literature is extensive and demonstrates that general conclusions regarding VOI are difficult to draw. However, the focus thus far has been decision situations involving binary decisions or decisions with only a few alternatives. In no case have general conclusions been investigated in resource allocation or knapsack problems. Furthermore, although the resource-allocation literature does include some discussion of the VOI in knapsack settings, the properties of information value have not been investigated.

3. Value of Information in Stochastic Knapsack Problems: Suppose that before choosing the optimal knapsack, a decision maker can purchase an information system Θ that yields a vector $\boldsymbol{\theta}' = (\theta_1,\ldots,\theta_N)$ of signals regarding the value of each item v_i. The value of the information system is, according to Equation (2), the b' that solves

$$E_{\boldsymbol{\theta}}\begin{bmatrix} \max_{\mathbf{x}} E_{\mathbf{v}|\boldsymbol{\theta}}[u(V(\tilde{\mathbf{v}}\,|\,\boldsymbol{\theta},\mathbf{x})-b')] \\ st \quad \mathbf{w}'\mathbf{x} \le C \\ x_i \in \Omega \end{bmatrix} = \begin{bmatrix} \max_{\mathbf{x}} E_{\mathbf{v}}[u(V(\tilde{\mathbf{v}},\mathbf{x}))] \\ st \quad \mathbf{w}'\mathbf{x} \le C \\ x_i \in \Omega \end{bmatrix}. (4)$$

The left-hand side of Equation (4) requires the solution of a knapsack problem for each possible information vector $\boldsymbol{\theta}$ and candidate value b, which could be computationally intensive. If the decision maker's utility function exhibits CARA, then the value of the information system can be found by taking the difference between the certain equivalent with a free information system and the certain equivalent without the information system, as in Equation (3). This condition is satisfied when the DM is risk neutral, in which case he/she is only concerned with the posterior mean $(\overline{\mathbf{v}}\,|\,\boldsymbol{\theta})$, and the value of the information system is

$$b' = E_{\boldsymbol{\theta}}\begin{bmatrix} \max_{\mathbf{x}}(\overline{\mathbf{v}}\,|\,\boldsymbol{\theta})'\mathbf{x} \\ st \quad \mathbf{w}'\mathbf{x} \le C \\ x_i \in \Omega \end{bmatrix} - \begin{bmatrix} \max_{\mathbf{x}} \overline{\mathbf{v}}'\mathbf{x} \\ st \quad \mathbf{w}'\mathbf{x} \le C \\ x_i \in \Omega \end{bmatrix}. (5)$$

4. The 0-1 SKP: The remainder of this paper focuses on the 0-1 SKP by letting $\Omega = \{0,1\}$ in Equation (4). We further assume that there is a set of N items whose uncertain values v_i are i.i.d. random variables. All items are of equal weight. Without loss of generality, we assume this weight is equal to 1. The DM faces a resource capacity constraint equal to C. Assume that before choosing the optimal knapsack, the risk-neutral DM is able to obtain an unbiased i.i.d. information

signal θ_i regarding the value of item i. Assume further that the true value of the item and the information signal are relevant.

By Equation (5), the VOI is

$$VOI = E_{\boldsymbol{\theta}}\underbrace{\begin{bmatrix} \max_{\mathbf{x}} \sum_{i=1}^{N}(\overline{v}_i\,|\,\theta_i)x_i \\ st \sum_{\iota=1}^{N} x_i \le C \\ x_i \in \{0,1\} \end{bmatrix}}_{\text{part I}} - \underbrace{\begin{bmatrix} \max_{\mathbf{x}} \sum_{i=1}^{N}\overline{v}_i x_i \\ st \sum_{\iota=1}^{N} x_i \le C \\ x_i \in \{0,1\} \end{bmatrix}}_{\text{part II}}, \quad (6)$$

where
N denotes the number of items,
v_i denotes the item value,
θ_i denotes the information signal,
C denotes the capacity, $0 < C \le N$, $C = 1, 2, 3\ldots$
$(\overline{v}_i\,|\,\theta_i)$ denotes $E_{\mathbf{v}}(v_i\,|\,\theta_i)$,
$\overline{v}_i$ denotes Ev_i.

Part I is a "wait-and-see problem," whose objective value must be at least as large as the objective value of Part II. Thus VOI is nonnegative. The model in the bracket of Part I is an SKP with cardinality constraint. This problem is also called the "selecting problem," since the optimum objective value is the sum of the top C item values.

Before proceeding, we describe the concept of order statistics and state a result that we will use later in the paper.

<u>Order Statistics</u>
Consider random variables X_1, …, X_N. The first (smallest) order statistic is

$$X_{(1:N)} = \min(X_1,\ldots,X_N).$$

The last (largest) order statistic is

$$X_{(N:N)} = \max(X_1,\ldots,X_N).$$

The rth order statistic is then denoted $X_{(r:N)}$ and defined as

$$X_{(r:N)} = r^{th} \text{ smallest } X_i, i = 1, 2, \ldots, N.$$

In this paper, we will make extensive use of the means of order statistics, which we define as

$$EX_{(r:N)} = \int x f_{(r:N)}(x)\,dx,$$

where $f_{(r:N)}(x)$ is the probability density function (pdf) of $X_{(r:N)}$.

To facilitate developments later in the paper, we provide two useful propositions regarding order statistics [32].

Proposition 1 (P1): If the pdf describing each random variable $X_1, \ldots, X_N$ is symmetric about $x = 0$, then

$$EX_{(r:N)} = -EX_{(N-r+1:N)} .$$

Proposition 2 (P2): For any arbitrary distribution with finite k^{th} moment,

$$(n-r)EX_{(r:n)} + rEX_{(r+1:n)} = nEX_{(r:n-1)} .$$

P1 states that if the underlying pdfs are symmetric about zero, possibly after transformation, then their order statistics are also symmetric about zero.

P2 provides an iterative formula to determine mean of higher order statistics from lower order ones.

Since we have assumed that v_i and θ_i are i.i.d and that all items have the same weight, we can write the optimum objective value of part I of Equation (6) as

$$E_\theta \left(\sum_{r=N-C+1}^{N} X_{(r:N)} \right) = \sum_{r=N-C+1}^{N} E_\theta X_{(r:N)} ,$$

where X denotes $(\bar{v} \mid \theta)$. That is, the optimal objective value is simply the sum of the top r order statistic means.

The optimum objective value of Part II of Equation (6) is simply $C\bar{v}$. So,

$$
\begin{aligned}
VOI &= \sum_{r=N-C+1}^{N} EX_{(r:N)} - C\bar{v} \\
&= \sum_{r=N-C+1}^{N} E(X - \bar{v})_{(r:N)} \qquad (7) \\
&= \sum_{r=N-C+1}^{N} \mu_{(r:N)} ,
\end{aligned}
$$

where μ denotes $E(X - \bar{v})$.

4.1 VOI Sensitivity to Capacity: In this section, we investigate the relationship between VOI and the capacity constraint. We show that contrary to intuition, a tighter resource constraint (lower C) does not always imply that less should also be spent on information gathering.

We now consider VOI to be a function of the knapsack capacity, *VOI(C)*, and state two theorems.

Theorem 1. For continuous random variables v and θ, if the pdf of $(\bar{v} \mid \theta)$ is symmetric about $\bar{v}$, the maximum of the function *VOI(C)* is obtained at $C = \lceil N/2 \rceil$ or $C = \lfloor N/2 \rfloor$.*

---

* $\lceil \ \rceil$ is the ceiling function, and $\lfloor \ \rfloor$ is the floor function.

Proof. $(\bar{v} \mid \theta) - \bar{v}$ is symmetrically distributed about $\theta = 0$, since the pdf of $(\bar{v} \mid \theta)$ is symmetric about $\bar{v}$. Thus, by P1,

$$\mu_{(r:N)} = -\mu_{(N-r+1:N)} \quad . \qquad (8)$$

By Equation (7),

$$VOI = \sum_{r=N-C+1}^{N} \mu_{(r:N)}.$$

If N is even, the optimum objective value is

$$VOI_{\max} = \sum_{r=N/2+1}^{N} \mu_{(r:N)} , \text{ when } C = N/2 .$$

If N is odd, then by Equation (8),

$$\mu_{((N+1)/2:N)} = 0 .$$

When $C = (N+1)/2$ or $C = (N-1)/2$ the optimum objective value is

$$VOI_{\max} = \sum_{r=(N+1)/2}^{N} \mu_{(r:N)} = \sum_{r=(N-1)/2}^{N} \mu_{(r:N)} . ,$$

Thus, the maximum of *VOI(C)* is obtained at $C = \lceil N/2 \rceil$ or $C = \lfloor N/2 \rfloor$. $\qquad \square$

Theorem 2. For continuous random variables v and θ, if the pdf of $(\bar{v} \mid \theta)$ is symmetric about $\bar{v}$, *VOI(C)* is symmetric about $C = N/2$.

Proof. Since, by P1,

$$\mu_{(r:N)} = -\mu_{(N-r+1:N)} ,$$

$$VOI(C) = \sum_{r=N-C+1}^{N} \mu_{(r:N)} = \sum_{r=C+1}^{N} \mu_{(r:N)} = VOI(N-C) .$$

So, $VOI(C) = VOI(N-C)$. $\qquad \square$

Theorems 1 and 2 reveal that VOI as a function of C reaches a maximum when the capacity constraint is equal to one-half of the number of items. The practical implication is that VOI may increase as the capacity constraint tightens, meaning that we should spend more resources to understand the item values. However, if the capacity constraint tightens beyond a threshold, then VOI may decrease.

4.2 VOI Sensitivity to the Number of Items: In this section, we investigate the relationship between VOI and the number of items under consideration in SKP. Intuitively, the greater the number of items that may be included in the KP, the greater the VOI.

We now focus on the maximum VOI as a function of the number of items n, which we represent by $VOI_{\max}(n)$, and state the following theorem.

Theorem 3. For continuous random variables v and θ, if the pdf of $(\overline{v}\,|\,\theta)$ is symmetric about $\overline{v}$, the function $VOI_{\max}(n)$ is nondecreasing in n.

Proof. We know

$$VOI_{\max}(n) = \sum_{r=\lceil (n+1)/2 \rceil}^{n} \mu_{(r:n)} \;,$$

and want to show

$$\sum_{r=\lceil (n+2)/2 \rceil}^{n+1} \mu_{(r:n+1)} \geq \sum_{r=\lceil (n+1)/2 \rceil}^{n} \mu_{(r:n)} \;,$$

where μ denotes $E(\overline{v}\,|\,\theta) - \overline{v}$. We prove this by induction.

By Equation (7) when $n = 1$, $\mu_{(2:2)} \geq \mu_{(1:1)}$. Assume that the inequality holds for $n = k$, where k is even. We then have

$$\mu_{((k+2)/2:k+1)} + \mu_{((k+2)/2+1:k+1)} + \ldots + \mu_{(k+1:k+1)} \geq$$
$$\mu_{(k/2:k)} + \mu_{(k/2+1:k)} + \ldots + \mu_{(k:k)}.$$

Now, when $n = k+1$, we have

$$\mu_{(k+1:k+2)} + (k+1)\mu_{(k+2:k+2)} = (k+2)\mu_{(k+1:k+1)},$$
$$2\mu_{(k:k+2)} + k\mu_{(k+1:k+2)} = (k+2)\mu_{(k:k+1)},$$
$$\ldots\ldots \qquad (9)$$
$$\left(\frac{k}{2}+1\right)\mu_{((k+2)/2:k+2)} + \left(\frac{k+2}{2}\right)\mu_{(\frac{k+2}{2}+1:k+2)} = (k+2)\mu_{(\frac{k}{2}+1:k+1)}.$$

Summing all the terms for each side of Equation (9) yields

$$\sum_{r=k/2+1}^{k+2} \mu_{(r:n+1)} = \frac{k+2}{k+1}\sum_{r=k/2+1}^{k+1}\mu_{(r:n)} + \frac{k/2}{k+1}\mu_{(k/2+1:k+2)}.$$

Since $\mu_{(k/2+1:k+2)} \geq 0$,

$$\sum_{r=k/2+1}^{k+2} \mu_{(r:k+2)} > \sum_{r=k/2+1}^{k+1} \mu_{(r:k+1)}\;.$$

If k is odd, we can show it in the same way. Thus,

$$\sum_{r=\lceil (n+2)/2 \rceil}^{n+1} \mu_{(r:n+1)} \geq \sum_{r=\lceil (n+1)/2 \rceil}^{n} \mu_{(r:n)}\;. \qquad \square$$

Thus, considering more items strictly increases VOI. Unfortunately, as shown in Appendix, $VOI_{\max}(n)$ may not be linear.

4.3 VOI Sensitivity to the Degree of Prior Uncertainty: In this section, we investigate the relationship between VOI and the degree of prior uncertainty regarding item values. We take the degree of uncertainty to be measured by the standard deviation of the item values. Intuitively, greater uncertainty should increase VOI. However, as has been shown in the case of single projects, this may be incorrect.

We further restrict the conditional expectation $(\overline{v}\,|\,\theta)$ to be a linear function of the standard deviation of the item values, such that $(\overline{v}\,|\,\theta) = m(\theta)\sigma_v + t$. This property would hold for jointly normal item values and signals, for example. Since $VOI(C)$ is symmetric, we focus on maximum VOI, which occurs at $C = \lceil N/2 \rceil$. We now state a somewhat surprising theorem regarding prior uncertainty.

Theorem 4. For continuous random variables v and θ, if the pdf of $(\overline{v}\,|\,\theta)$ is symmetric about $\overline{v}$, and the conditional expectation $(\overline{v}\,|\,\theta)$ is a linear function of the standard deviation σ_v of v, the function $VOI(\sigma_v)$ is linearly increasing.

Proof. Since $(\overline{v}\,|\,\theta) = m(\theta)\sigma_v + t$, $(\overline{v}\,|\,\theta) - \overline{v}$ is also a linear function of σ_v, say $(\overline{v}\,|\,\theta) - \overline{v} = m(\theta)\sigma_v + t'$.

$$VOI(\sigma_v) = \sum_{r=N-C+1}^{N} \mu_{(r:N)} = \sum_{r=N-C+1}^{N} E\left((\overline{v}\,|\,\theta) - \overline{v}\right)_{(r:N)}$$
$$= \sum_{r=N-C+1}^{N} \left(\sigma_v Em(\theta)_{(r:N)} + t'\right)$$
$$= \sigma_v \sum_{r=N-C+1}^{N} Em(\theta)_{(r:N)} + \lceil N/2 \rceil t'$$

Thus, the $VOI(\sigma_v)$ is linearly increasing with respect to prior uncertainty regarding item values. $\square$

Therefore, as our prior uncertainty increases, we should spend (strictly) more on information gathering. Although obtained under rather strict assumptions, the result is surprising given the negative results stated in §3.

4.4 VOI Sensitivity to Information System Quality: In this section, we investigate the sensitivity of VOI to the quality of the information system, which we take to be measured by the correlation coefficient ρ. We maintain our assumption that the item-value conditional means are a linear function of the correlation coefficient. A correlation coefficient of 1 would represent perfect information.

We further restrict that the conditional expectation $(\bar{v}\,|\,\theta)$ is a linear function of the correlation between the item values and signals, say $(\bar{v}\,|\,\theta) = m'(\theta)\rho + t''$, where m' is a function of random variable θ and t'' is a scalar, without loss of generality. We now state our final theorem.

Theorem 5. For continuous random variables v and θ, if the pdf of $(\bar{v}\,|\,\theta)$ is symmetric about $\bar{v}$, and the conditional expectation $(\bar{v}\,|\,\theta)$ is a linear function of the correlation ρ between v and θ, the function $VOI(\rho)$ is linearly increasing.

Proof. This proof is similar to that of Theorem 4. Because $(\bar{v}\,|\,\theta) = m'(\theta)\rho + t''$, $(\bar{v}\,|\,\theta) - \bar{v}$ is also a linear function of σ_v, say $(\bar{v}\,|\,\theta) - \bar{v} = m'(\theta)\rho + t'''$.

$$
\begin{aligned}
VOI(\rho) &= \sum_{r=N-C+1}^{N} \mu_{(r:N)} = \sum_{r=N-C+1}^{N} E\big((\bar{v}\,|\,\theta) - \bar{v}\big)_{(r:N)} \\
&= \sum_{r=N-C+1}^{N} \big(\rho Em(\theta)_{(r:N)} + t'''\big) \\
&= \rho \sum_{r=N-C+1}^{N} Em(\theta)_{(r:N)} + \lceil N/2 \rceil t'''
\end{aligned}
$$

Thus, $VOI(\rho)$ is linearly increasing with respect to information system quality. □

5. Examples: In this section, we demonstrate the properties proved in the previous sections, by applying Monte Carlo simulation with 1000 iterations. After each iteration, we solve the first part of Equation (6) using a dynamic programming algorithm [27].

To match the assumptions made in §4, we consider a 0-1 SKP whose uncertain values v_i and signals θ_i are i.i.d and jointly normal. The weight of each item is assumed to equal 1. These assumptions satisfy those made in §4. For example, under joint normality, the conditional value of the item values is

$$
(\bar{v}\,|\,\theta) = \bar{v} + \rho\frac{\sigma_v}{\sigma_\theta}\big(\theta - \bar{\theta}\big),
$$

whose pdf is symmetric about $\bar{v}$.

We consider the specific case of 10 items whose means and standard deviations are 10 and 2.5, respectively.

5.1 Simulation-VOI Sensitivity to Capacity: Two cases are simulated. In the first, the coefficient between item values and signals is 0.8. In the second, it is 0.2.

The simulation results are displayed in Figure 1. As proved in Theorems 1 and 2, $VOI(C)$ is maximized at and symmetric about $C = 5$.
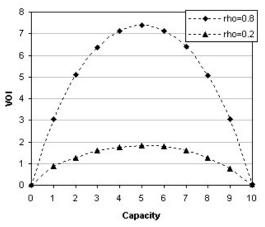

Figure 2: Simulation results of VOI as a function of capacity

5.2 Simulation-VOI Sensitivity to the Number of Items: We simulate two cases, with correlation coefficients of 0.2 and 0.8, respectively. As in Theorem 3, the capacity is fixed at $C = \lceil N/2 \rceil$. The simulation results appear in Figure 3 and demonstrate that $VOI_{max}(n)$ is nondecreasing but not linear.
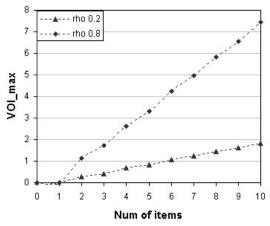

Figure 3: Simulation results of VOI as a function of number of items

5.3 Simulation-VOI Sensitivity to the Degree of Prior Uncertainty: In this case, we simulate $VOI_{max}(\sigma_v)$ for standard deviations ranging from 0 to 10. As before, we consider correlation coefficients of 0.2 and 0.8. The results are displayed in Figure 4.
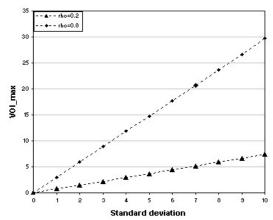
Figure 4: Simulation results of VOI as a function of prior uncertainty

The simulation results demonstrate that $VOI_{max}(\sigma_v)$ is increasing linearly with respect to prior uncertainty of item values and that *VOI* is equal to 0 when the standard deviation is 0.

5.4 Simulation-Sensitivity to Information System Quality: In this section, we simulate *VOI* as a function of the correlation coefficient. The results are displayed in Figure 5.
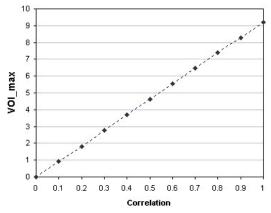

Figure 5: Simulation results of VOI as a function of information-system quality

As proved in Theorem 5, *VOI* is linearly increasing in the correlation coefficient.

6. Conclusions and Future Work: This research has explored the properties of information value within the context of resource-allocation decisions made under uncertainty. We have also shown that when resources are constrained, it may be better to spend more on information gathering—tighter budget constraints in one area should be met with greater budgets in another. We have also proved that some VOI properties that do not hold in the case of single projects do hold in the

knapsack setting. For example, we have demonstrated that VOI is strictly increasing in the prior uncertainty.

We are still investigating several additional problems within the context of the 0-1 SKP. These include the properties of VOI as a function of the decision maker's risk attitude, the impact of dependence between item values, dependence between signals, and the relationship between perfect and imperfect information.

Once these investigations are completed, we will attempt to extend our results to unbounded, bounded, multidimensional, and multiple SKPs.

8. Appendix:

We demonstrate that VOI may not be a linearly increasing function of the number of items.

We assume item values v_i are i.i.d. normal random variables. The true value of the item and the information signal θ_i are jointly normally distributed with a correlation ρ. $(\bar{v} \mid \theta)$ is $\bar{v} + \rho \frac{\sigma_v}{\sigma_\theta}(\theta - \bar{\theta})$, whose pdf is symmetric about $\theta = \bar{v} = 10$.

When $n = 1$,
$$VOI_{max}^n(1) = 0.$$
When $n = 2$,
$$VOI_{max}^n(2) = Ev - \frac{\rho\sigma_v}{\sigma_\theta}E\theta + \frac{\rho\sigma_v}{\sigma_\theta}E\theta_{(2:2)}.$$
When $n = 3$,
$$VOI_{max}^n(3) = 2Ev - 2\frac{\rho\sigma_v}{\sigma_\theta}E\theta + \frac{\rho\sigma_v}{\sigma_\theta}\left(E\theta_{(2:3)} + E\theta_{(3:3)}\right).$$
Let Δ_i denote $VOI_{max}^n(i+1) - VOI_{max}^n(i)$.
$$\Delta_1 = Ev - \frac{\rho\sigma_v}{\sigma_\theta}E\theta + \frac{\rho\sigma_v}{\sigma_\theta}E\theta_{(2:2)},$$
$$\Delta_2 = Ev - \frac{\rho\sigma_v}{\sigma_\theta}E\theta + \frac{\rho\sigma_v}{\sigma_\theta}\left(E\theta_{(2:3)} + E\theta_{(3:3)} - E\theta_{(2:2)}\right).$$
If the function $VOI_{max}^n(n)$ is linear, then $\Delta_1 = \Delta_2$. Thus,
$$E\theta_{2:2} = E\theta_{(2:3)} + E\theta_{(3:3)} - E\theta_{(2:2)}. \qquad (10)$$
By Proposition 2,
$$E\theta_{(2:3)} + 2E\theta_{(3:3)} = 3E\theta_{(2:2)}. \qquad (11)$$
Substituting Equation (10) into Equation (11) yields $E\theta_{(3:3)} = E\theta_{(2:2)}$, which is a contradiction. Thus, the function $VOI_{max}^n(n)$ may not be linear.

9. References:

[1] United States, Department of Homeland Security. 2006. Budget-in-brief: Fiscal Year 2007, Washington, DC.

[2] United States, Department of Homeland Security. 2006. National infrastructure protection plan, Washington, DC.

[3] United States, Presidential/Congressional Commission on Risk Assessment and Risk Management. 1997. Risk assessment and risk management in regulatory decision-making, Washington, DC.

[4] Morita, Hiroshi, Hiroaki Ishii and Toshio Nishida. 1989. Stochastic linear knapsack programming problem and its application to a portfolio selection problem. *European Journal of Operational Research* **40** 329-336.

[5] Henig, Mordechai I. 1990. Risk criteria in a stochastic knapsack problem. *Operations Research* **38**(5) 820-825.

[6] Morton, David P. and R. Kevin Wood. 1998. On a stochastic knapsack problem and generalizations. In: D.L. Woodruff (ed.), *Advances in computational and stochastic optimization, logic programming and heuristic search: Interfaces in computer science and operations research*. Kluwer Academic Publishers, Norwell, Massachusetts, pp. 149-168.

[7] Steinberg, E. and M. S. Parks. 1979. A preference order dynamic program for a knapsack problem with stochastic rewards. *The Journal of the Operational Research Society* **30**(2) 141-147.

[8] Sniedovich, Moshe. 1980. Preference order stochastic knapsack problems: Methodological issues. *The Journal of the Operational Research Society* **31**(11) 1025-1032.

[9] Carraway, Robert L., Robert L. Schmidt and Lawrence R. Weatherford. 1993. An algorithm for maximizing target achievement in the stochastic knapsack problem with normal returns. *Naval Research Logistics* **40** 161-173.

[10] von Neumann, J. and O. Morgenstern. 1947. *Theory of games and economic behavior*, Second ed. Princeton University Press, Princeton, NJ.

[11] Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber and David Heath. 2002. Coherent measures of risk. *Mathematical Finance* **9**(3) 203-228.

[12] Hilton, Ronald W. 1981. The determinants of information value: Synthesizing some general results. *Management Science* **27**(1) 57-64.

[13] Mehrez, Abraham. 1985. The effect of risk aversion on the expected value of perfect information. *Operations Research* **33**(2) 455-458.

[14] LaValle, Irving H. 1968. On cash equivalents and information evaluation in decisions under uncertainty: Part iii: Exchanging partition-j for partition-k information. *Journal of the American Statistical Association* **63**(321) 285-290.

[15] LaValle, Irving H. 1968. On cash equivalents and information evaluation in decisions under uncertainty: Part i: Basic theory. *Journal of the American Statistical Association* **63**(321) 252-276.

[16] LaValle, Irving H. 1968. On cash equivalents and information evaluation in decisions under uncertainty: Part ii: Incremental information decisions. *Journal of the American Statistical Association* **63**(321) 277-284.

[17] Morris, James R. 1974. The logarithmic investor's decision to acquire costly information. *Management Science* **21**(4) 383-391.

[18] Gould, John P. 1974. Risk, stochastic preference, and the value of information. *Journal of Economic Theory* **8**(64) 64-84.

[19] Mehrez, Avraham and Alan Stulman. 1982. Some aspects of the distributional properties of the expected value of perfect information (evpi). *The Journal of the Operational Research Society* **33**(9) 827-836.

[20] Fatti, L. P. 1987. Bounds and properties of the expected value of sample information for a project-selection problem. *Naval Research Logistics* **34** 141-150.

[21] Blackwell, David. 1953. Equivalent comparisons of experiments. *The Annals of Mathematical Statistics* **24**(2) 265-272.

[22] Clemen, Robert T. and Robert L. Winkler. 1985. Limits for the precision and value of information from dependent sources. *Operations Research* **33**(2) 427-442.

[23] Samson, D., A. Wirth and J. Rickard. 1989. The value of information from multiple sources of uncertainty in decision analysis. *European Journal of Operational Research* **39** 254-260.

[24] Keisler, Jeffrey M. 2005. Additivity of information value in two-act linear loss decisions with normal priors. *Risk Analysis* **25**(2) 351-359.

[25] Bickel, J. Eric. 2008. The relationship between perfect and imperfect information in a two-action risk-sensitive problem. *Decision Analysis (Articles in Advance)* 1-13.

[26] Martello, Silvano and Paolo Toth. 1990. *Knapsack problems: Algorithms and computer implementations*. John Wiley & Sons Ltd., West Sussex, England.

[27] Kellerer, Hans, Ulrich Pferschy and David Pisinger. 2005. *Knapsack problems*. Springer-Verlag, Heidelberg.

[28] Mehrez, Abraham and Alan Stulman. 1984. A note on capital budgeting in the purchase of information. *Managerial and Decision Economics* **5**(4) 251-254.

[29] Mehrez, Abraham and Awanti P. Sethi. 1989. Hierarchical planning of project-selection problems

with information purchasing. *The Journal of the Operational Research Society* **40**(3) 267-279.

[30] Keisler, Jeffrey M. 2004. Value of information in portfolio decision analysis. *Decision Analysis* **1**(3) 177-189.

[31] Bickel, J. Eric, Richard L. Gibson, Duane A. McVay, Stephen Pickering and John Waggoner. 2006. Quantifying 3d land seismic reliability and value. *SPE Annual Technical Conference and Exhibition.* Paper **102340**.

[32] David, H. A. and H. N. Nagaraja. 2003. *Order statistics*, Third ed. John Wiley & Sons, Inc., Hoboken, New Jersey.