October 2004

Technical note: Comparative static analysis of information value in a canonical decision problem

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

TECHNICAL NOTE: COMPARATIVE STATIC ANALYSIS OF INFORMATION VALUE IN A CANONICAL DECISION PROBLEM.

JEFFREY M. KEISLER
University of Massachusetts-Boston

ABSTRACT

To gain insight into the behavior of the value of information, this paper identifies specific rules for a canonical decision problem: the two-act linear loss decision with normal prior probability distributions. Conditions are derived for which the expected value of perfect information increases when mean and standard deviation are both linear functions of an exogenous variable. A variety of richer decision problems can be adapted to the problem, so that the general results obtained here can be immediately applied to understand drivers of information value.

INTRODUCTION

This paper considers the comparative statics of the expected value of perfect information in the two-act linear loss (TALL) decision problem with normal prior probability distributions when the mean and standard deviation depend on a common exogenous factor.

In this problem, the decision maker is an expected monetary value maximizer who must choose between the status quo and a new alternative. The increment in value from the new alternative compared to the status quo is an uncertain quantity, *$x*, where x is normally distributed with prior mean $\mu \geq 0$ and prior standard deviation σ. This is a canonical problem in the decision analysis literature (Raiffa & Schlaifer, 1961) and the closely related unit normal loss function is now prevalent enough to be a standard feature in major statistics and optimization software applications. It is well known and clearly seen that as μ increases, value of information decreases, and as σ increases, so does value of information.

If μ and σ can be expressed as functions of an exogenous variable k, $\mu(k)$ and $\sigma(k)$, then value of information for the decision can also be expressed as a function of k. This paper identifies conditions for under which the value of information is increasing or decreasing in such an exogenous variable. Of most interest are situations where both μ and σ are increasing in *k*. When the value of the new alternative is the product of the exogenous variable and some uncertain quantity, the relationships between *k* and μ and between *k* and σ are often linear. In this particular case, comparative static results can be expressed in terms of the parameters of these functional relationships. This analysis will facilitate future work on richer decision problems in which more complicated phenomena are modeled so as to make the functions $\mu(k)$ and $\sigma(k)$ tractable. After deriving these conditions, we shall consider in detail one illustrative example. Other applications along these lines are then proposed, where the exogenous factors are more complex than a scalar variable, but the linear formulation still applies.

BACKGROUND

The starting point for analyzing value of information in the TALL problem is Raiffa and Schlaifer's derivation of the complete analytical form for determining EVPI in any specific instance of the TALL problem with normal priors. The analytics of EVPI in individual decision making have received only intermittent attention since that time. Demski (1972) and later Hilton (1981) showed that few general conclusions hold concerning the value of information. Thon and Thorlund-Petersen (1993) consider the effect of risk and risk attitudes on EVPI. Closer to the current problem, Herath and Park (2001) examine an example illustrating the information value resulting from changes in mean and variance after partial resolution of uncertainty (in the same linear loss problem with normal distributions), this in the course of explaining the interesting parallel between value of information and financial option pricing.

Felli and Hazen (1998) used Monte Carlo techniques to examine the sensitivity of EVPI in decision problems to chance node probabilities and variability of decision parameters. They note that currently available Monte Carlo techniques allow for estimation of EVPI even when there are multiple interacting variables, and they caution that "EVPI values can be calculated in closed form only for problems with very simple or special structure. For most realistic decision problems, EVPI values must be numerically approximated." Brennan *et al*. (2002) illustrate the state of the art in simulation-based estimation of EVPI.

In contrast to these approaches, in this paper, our purpose is not to so much to calculate EVPI as to gain insight into the phenomenon of EVPI by studying a simple situation in closed form. As noted above, this constrains us to using a simple structure. For example, along with the other assumptions of the TALL problem, our analysis is generally restricted to EVPI with respect to one uncertainty at a time. In return, we can identify precise comparative static characteristics of EVPI in this class of decisions in the hopes that these qualities apply not only to the exact decision structures analyzed, but also to a wider range of situations that are more complicated and cannot be modeled so easily or precisely. These results could be complementary to Monte Carlo approaches, for example, suggesting patterns that one would look for and questions one might ask.

ANALYSIS

Notation:

V : EVPI for the decision.

f_{N*} : the probability density function for the standard unit normal distribution,

G_{N*} : the unit normal right tail cumulative probability function,

$u = \mu/\sigma$,

L_{N*} : Linear loss integral for the standard unit normal distribution.

H : The normal hazard function (= f/G), sometimes called the "inverse Mills ratio."

We shall assume throughout this analysis that $\sigma > 0$ and $\mu \geq 0$; for $\mu \leq 0$ the results would be similar with appropriate changes of sign. All derivatives should be interpreted as right-hand derivatives in order for results to include the case where $\mu = 0$. For $\sigma = 1$ and $\mu \geq 0$, Raiffa and Schlaifer derived the equation

$$V = f_{N*}(u) - uG_{N*}(u) \qquad (1)$$

also called the unit normal linear loss integral evaluated at u and denoted $L_{N*}(u)$; it represents the expected amount of loss avoided by having the option to take a certain 0 if the actual value of x turns out to be less than 0. When $\sigma \neq 1$ and $\mu \geq 0$ the value of information is equal to $\sigma L_{N*}(u)$:

$$V = \sigma[f_{N*}(u) - uG_{N*}(u)]. \qquad (2)$$

We now suppress N^* from the subscripts of G and f, and from u, because the rest of the paper will assume unit normal prior distributions. The term $f(u)$ denotes the probability density function evaluated at u, and $G(u)$ is the right tail cumulative density evaluated at u.

Several points are easily observed, starting with a formal statement of the well-known facts mentioned above.

1) $\boldsymbol{\partial V/\partial \mu = -G(u)}$. As the probability density function shifts to the right, the values of the alternatives grow farther apart, the region representing loss avoided (to the left of 0) shrinks, and the value of information decreases.
2) $\boldsymbol{\partial V/\partial \sigma = f(u)}$. As the probability density function spreads out, 0 is at a relatively higher part of the curve, and the curve extends farther out, so the value of information increases. If both the prior mean increases and the standard deviation decreases, the value of information decreases.

A third straightforward observation is that if the mean and the standard deviation are increased by a constant proportion, the value of information increases by the same proportion (because u is unchanged, while σ increases). It is less clear, however, what happens when the prior mean and variance both increase or both decrease, but by different proportions. This is an important question, because there are often factors that influence both μ and σ in the same direction, e.g. correlated risk and return. To answer this question, it is no longer enough merely to note that both effects exist. Rather, a method is needed to compare the size of the shifting and spreading effects. Proposition 1 derives a condition, featuring the hazard function, for determining which effect dominates.

Proposition 1: **Where $\mu \geq 0$ and $\sigma > 0$ are differentiable and monotonically increasing functions of an exogenous variable $k \geq 0$, the value of information is increasing in k for those values and only those values of k satisfying $H[\mu(k)/\sigma(k)] > (d\mu/dk)/(d\sigma/dk)$.**

Proof: The full derivative of V with respect to k is

$$(\partial V/\partial \mu)\,(d\mu/dk) + (\partial V/\partial \sigma)(d\sigma/dk) \tag{3}$$

$$= \quad -G(u)d\mu/dk + f(u)d\sigma/dk, \tag{4}$$

which is positive if and only if the inequality in Proposition 1 holds.

This condition requires only the estimation of two natural ratios: the first ratio is the common $\mu/\sigma = u$, from which $H(u)$ is easily computed[1]. The second ratio is $(d\mu/dk)/(d\sigma/dk)$, i.e., the degree to which the mean is more sensitive to k than is the standard deviation. Intuitively, the result explains which effect dominates. When $f(u)/G(u)$ is small, expected loss is more sensitive to the probability of loss than to the magnitude of loss. In this case, increasing μ, which is small, decreases the probability of error and has a large effect on V. When $f(u)/G(u)$ is large, mistakes are minor, so an increase in σ increases the magnitude of the mistakes and has a large effect on V. The latter case is more like insurance or R & D decisions, the former is more like a repetitive product launch decision or a process adoption decision.

With a stronger assumption about the dependence of μ and σ on k, we can be more specific, as indicated in Corollary 1.

For all the following results, we assume that there is a first order linear relationship between k and μ and between k and σ, specifically,

$\mu = ak + b$,

and

$\sigma = ck + d$,

where $a, c > 0$ and $k > -d/c$ (so that $\sigma > 0$).

***Corollary 1*: *V* is increasing in *k* for those values and only those values of *k* satisfying *H(u) > a/c*.**

---

[1] The hazard function is most commonly seen in reliability analysis. Intuitively, it represents the likelihood that x will be close to u if it is known that $x \geq u$. The normal hazard function approaches 0 from above as u goes to $-\infty$, is equal to $\sqrt{2/\pi}$ when u is 0, and approaches u from above as u goes to $+\infty$.

Proof:

In the key condition $H[\mu(k)/\sigma(k)] > (d\mu/dk)/(d\sigma/dk)$ of Proposition 1, substitute $u = \mu(k)/\sigma(k)$, $d\mu/dk = a$, $d\sigma/dk = c$ to obtain $H(u) > a/c$. Therefore by Proposition 1, value of information is increasing at k for those k and only those k satisfying the key condition $H(u) > a/c$.

In order to consider the question of when this condition implies that V(k) is increasing in k, we first note the following properties of H:

1. $H(u)$ is a positive increasing function of u which approaches 0 as $u \to -\infty$.
2. $H(u) > u$ and approaches u asymptotically as $u \to \infty$.
3. $H(0) = \sqrt{2/\pi}$.

We now consider two possibilities.

Case 1: $b/a > d/c$. **Conclusion:** $V(k)$ is always increasing in k.

The reasoning is as follows: In this case the function $u = \dfrac{ak+b}{ck+d}$ is a decreasing positive function of k in the feasible range $k > -d/c$, and u approaches a/c from above as $k \to \infty$. Then the key condition $H(u) > a/c$ of Corollary 1 holds automatically because $u > a/c$ for all feasible k, and $H(u) > u$ for all u. Therefore $V(k)$ is always increasing in k.

When d is non-zero, an equivalent condition is $b/d > a/c$. This situation is shown in FIGURE 1.

INSERT FIGURE 1 ABOUT HERE

Case 2: $b/a < d/c$. **Conclusion:** $V(k)$ is increasing for $k > k_0$, decreasing for $k < k_0$, and therefore reaches a minimum at $k = k_0$, where $k_0 = -\dfrac{du_0 - b}{cu_0 - a}$, and $u_0 = H^{-1}(a/c)$.

The reasoning is as follows: In this case, the function $u = \dfrac{ak+b}{ck+d}$ is an increasing function of k in the feasible range $k > -d/c$, and u approaches a/c from below as $k \to \infty$. The key condition $H(u) > a/c$ of Corollary 1 holds for $u > H^{-1}(a/c) \equiv u_0$. Because $u_0 < H(u_0) = a/c$, and u approaches a/c from below, the increasing function u of k will exceed u_0 for k large enough, specifically, for k exceeding $k_0 \equiv -\dfrac{du_0 - b}{cu_0 - a}$.

So the key condition $H(u) > a/c$ holds for $k > k_0$, and the opposite condition holds for $k < k_0$. We conclude that $V(k)$ is increasing for $k > k_0$ and decreasing for $k < k_0$, and therefore reaches a minimum at $k = k_0$.

This situation is shown in FIGURE 2.

INSERT FIGURE 2 ABOUT HERE

If we add the constraints $k, b \geq 0, d > 0$, Corollary 1 and the properties of H immediately imply that V is increasing in k if $a/c < \sqrt{(2/\pi)}$ (because then $H(u) > \sqrt{(2/\pi)} > a/c$), or if $a/c > \sqrt{(2/\pi)}$ and $H(b/d) > a/c$ (because $H(u)$ lies between $H(b/d)$ and $H(a/c)$).

These results tell whether EVPI is increasing, decreasing, or has a minimum in an exogenous variable of which mean and standard deviation are linear functions. With these results, we now explore richer specifications of the relationship between exogenous variables and summary statistics.

Illustrative example

In the following illustrative example, the analytic results are useful. The physical decision is described, followed by the information gathering decision. Consider this situation: A utility company has the opportunity to build a plant to serve a region. The generation cost per kilowatt hour (kWh) is, for a known process, directly proportional to the unit cost for fuel. Fuel costs are uncertain. The price per kWh is fixed by contract with the regional government, and it is set high enough that the company can expect to cover the cost of building the plant and make some profit serving the region. The plant has the additional capability to serve a second region, and would be required to charge the same price there. The quantity demanded by the core region and other regions is known.

Before signing the contract, the company would like to resolve uncertainty about its fuel costs. It is willing to pay a research firm a fee to resolve the uncertainty.

In algebraic terms, the problem is as follows:

$$\text{Profit} = (P\text{-}C)\ (Q_1 + Q_2) - F, \tag{5}$$

where P is price per kWh, C is cost per kWh, Q_1 and Q_2 are quantity of demand in regions 1 and 2, and F is the fixed cost. Prior beliefs about C are represented as a normal distribution with mean μ_C and standard deviation σ_C. The value of information about cost can be computed noting that the decision of whether or not to build the plant is a TALL decision. The mean value (μ) of the non-zero alternative is

$$(P\text{-}\ \mu_C)\ (Q_1 + Q_2) - F, \tag{6}$$

which is the profit in (5), substituting μ_C for C, and its standard deviation (σ) is equal to $\sigma_C(Q_1 + Q_2)$. Again, letting $u = \mu/\sigma$, the value of information is $\sigma[f(u) - uG(u)]$.

The company then receives news that the situation has changed. Should it now be willing to pay more or less to resolve uncertainty about cost prior to deciding whether to build the plant? The answer, of course, depends on what the change was.

Any of the parameters, P, F, μ_C, σ_C, Q_1 and Q_2 could have increased or decreased. Using the results in this paper, it is possible to answer the question about value of information without calculating value of information. If P or F increases or if μ_C decreases, μ increases and there is no effect on σ, so the value of information decreases. If σ_C increases, σ increases and there is no effect on μ so the value of information increases. Q_1 cannot change without changing the other terms of the contract, because the price is set so as to cover fixed costs, which may conflict with the story line.

The interesting case is when Q_2 increases. In this case it is unclear whether value of information about cost increases or decreases, because both the standard deviation and the gap between the two alternatives increase. To apply the main analytic results, we assign Q_2 as the exogenous variable, k. Then,

$$a = P\text{-}\ \mu_C, \tag{7}$$

$$b = (P\text{-}\ \mu_C)Q_1 - F, \tag{8}$$

$c = \sigma_C$ and $d = Q_1\sigma_C$. In order determine whether the value of information about C is increasing in Q_2, we first compute

$$a/c = (P - \mu_C)/\sigma_C \tag{9}$$

and

$$b/d = [(P - \mu_C)(Q_1 - F)] / (Q_1 \times \sigma_C). \tag{10}$$

This is illustrated by placing numbers in the example, and observing the results for each of the cases described in the previous section. In each case, calculations are continued in order of difficulty until one of the conditions is met to indicate the direction of change in V. The predicted change is then compared with the actual result when Q_2 is incremented by 1.

Example 1: $P = 34$, $\mu_C = 30$, $F = -5$, $Q_1 = 5$, $Q_2 = 5$, $\sigma_C = 3$.

b/a $(= 6.25) > d/c$ $(= 5)$, so $b/a > d/c$. Case 1 applies.

Prediction: V is increasing in Q_2.

Actual result: $V(5) = 0.879$, $V(6) = 1.001$.

Example 2: $P = 34$, $\mu_C = 30$, $F = 15$, $Q_1 = 5$, $Q_2 = 5$, $\sigma_C = 2$.

$k_0 = 12.517 > k$. Corollary 1 case 2 applies.

Prediction: V is decreasing in Q_2 reaching a minimum at $k_0 = 12.517$.

Actual result: $V(5) = 1.012$, $V(6) = 0.964$.

Note that $V(k_0) = 0.870$ and that at this point, $H(u) = 2 = a/c$.

Example 3: $P = 32$, $\mu_C = 30$, $F = 10$, $Q_1 = 10$, $Q_2 = 5$, $\sigma_C = 3$.

If the constraints $k, b \geq 0, d > 0$ are known to apply, note that $a/c = 0.667 < \sqrt{(2/\pi)}$ so $k_0 < 0$.

If these constraints do not apply, compute $k_0 = -6.216 < k$. Corollary 1, case 2 applies.

Prediction: V is increasing in Q_2.

Actual Result: $V(5) = 9.697$, $V(6) = 10.126$.

Example 4: $P = 34$, $\mu_C = 30$, $F = 10$, $Q_1 = 10$, $Q_2 = 5$, $\sigma_C = 3$.

If the constraints $k, b \geq 0, d > 0$ are known to apply, $H(b/d) = 1.525 > a/c$.

If these constraints do not apply, compute $k_0 = -4.227 < k$. Corollary 1, case 2 applies.

Prediction: V is increasing in Q_2.

Actual result: $V(5) = 3.021$, $V(6) = 3.134$.

Thus, it is not necessary to perform value of information calculations in order to recognize whether changes in the exogenous variable (Q_2) are favorable or unfavorable to the case for acquiring information.

OTHER APPLICATIONS

It is possible to model other decision problems so as to make their information sources fit this structure, enabling analysis of the rich set of situations from which normal distributions arise. This was, in fact, the impetus for this paper. Specifically, normally distributed variables remain normal under multiplication by constants and addition of constants or of other normally distributed variables. These facts allow a variety of decisions to be reduced to the current problem, for example, when $X_i \sim N[\mu_i(k), \sigma_i(k)]$ and $X = \Sigma(X_i)$. This could arise from a case where X is the total profit from a product, and X_i is the profit from the product in each of n regions. In particular, if $\mu_i = a_i k + b_i$ and $\sigma_i = c_i k$ and the X_i are uncorrelated, the prior distribution on X is normal with mean $\mu(k) = k\Sigma\, a_i + \Sigma\, b_i$ and standard deviation $\sigma(k) = k\ \sqrt{\Sigma\, c_i^2}$.

A related example is where $X: <(p_1,X_1),(p_2,X_2), \ldots>$, where X_i is the value and p_i is the probability for the ith branch of a chance node (so $\Sigma\, p_i = 1$) and $X_i \sim N(\mu_i\, k, \sigma_i k)$, where uncertainty about the X_i can be resolved but the chance node cannot be resolved prior to choosing this path. The expected value of X has mean $\mu(k) = k\Sigma(p_i\mu_i)$ and, with respect to resolution of the uncertainty around X_i, X has standard deviation $\sigma(k) = k\ \sigma\sqrt{\Sigma(\sigma_i^2 p_i^2)}$.

In cases where $\mu(k)$ and $\sigma(k)$ are differentiable but not linear, the results may still be informative when interpreted with regard to the first-order Taylor expansion $\mu(k) = ak + b$, and the first order Taylor expansion of $\sigma(k) = ck + d$.

CONCLUSION

The comparative static results clarify what it means for a decision to become more clear-cut or more of a toss-up with respect to information about specific variables. In particular, if we know how exogenous variables affect both the mean and the standard deviation of the value of an alternative, the results in section 2 indicate which effect dominates – the increase in EVPI due to increased spread of the uncertain value, or the decrease in EVPI due to the increased gap between the expected value of the new alternative and the status quo. Where mean and standard deviation are linear functions of an exogenous variable, we can determine whether EVPI is increasing or decreasing in that variable by comparing several simple ratios involving the rates of change of mean and standard deviation, the hazard function of both of these, and sometimes the quantity $\sqrt{(2/\pi)}$. These results provide a generic map of "value of information space" for a significant class of decisions. To the extent that actual decisions resemble the two-act linear loss problem, the results predict the direction of change in information value (whether or not it has been explicitly calculated) when conditions change.

ACKNOWLEDGEMENTS

The author would like to thank Howard Raiffa, Terry Reilly, George Wu for their encouragement and advice, and also the anonymous referees for their helpful comments that improved this paper.

REFERENCES

[1] BRENNAN, ALAN, JIM CHILCOTT, SAMER KHARROUBI and TONY O'HAGAN, "A Two Level Monte Carlo Approach to Calculating Expected Value of Perfect Information – Resolution of the Uncertainty in Methods," working paper, University of Sheffield, England, 2002. (Retrieved July 2003 from http://www.shef.ac.uk/chebs/Papers/Brennan.pdf)

[2] DEMSKI, JOEL, *Information analysis*, Reading, MA, Addison Wesley, 1972.

[3] FELLI, JAMES C.. and GORDON B. HAZEN, "Sensitivity Analysis and the Expected Value of Perfect Information," *Medical Decision Making*, Vol. 18, No. 1, 1998, PP. 95-109.

[4] HERATH, HEMANTHA S.B. and CHAN S. PARK., "Real Options Valuation and its Relationship to Bayesian Decision-Making Methods," *The Engineering Economist*, Vol. 46, No. 1, 2001, pp. 1-32.

[5] HILTON, RONALD, "The Determinants of Information Value: Synthesizing Some General Results," *Management Science*, Vol. 27, No. 1, 1981, pp. 57-65.

[6] RAIFFA, HOWARD and ROBERT SCHLAIFER, *Applied Statistical Decision Theory*, Boston, Harvard Business School Division of Research, 1961.

[7] THON, DOMINIQUE and LARS THORLUND-PETERSEN, "The Value of Perfect Information in a Simple Investment Problem," *Information Economics and Policy*, Vol. 5, 1993, pp. 51-71.

Figure 1:V(k) is increasing in k because u(k) is always greater than a/c.

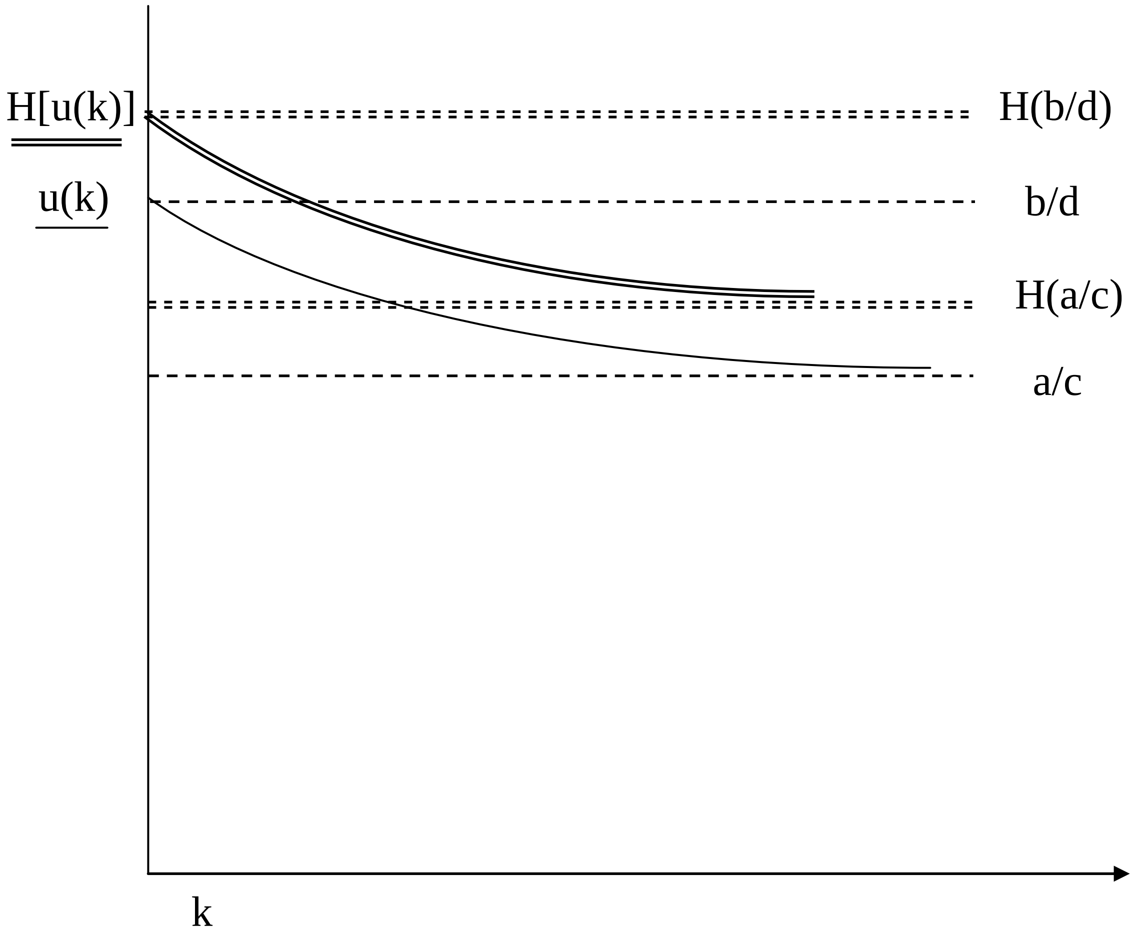

Figure 2: V(k) has a minimum in k where H[u(k)] = a/c.