

Value of Information Analysis Using Monte Carlo Simulations

Lutz Göhring

March 16, 2015

Executive Summary

Decision making processes can be categorized in two levels of decision making. 1. The actual problem of interest of a policy maker which we call the underlying decision on how to influence an ecological-economic system based on a particular information on the system available to the decision maker and 2. the meta decision on how to allocate resources to reduce uncertainty, i.e to increase the current information to improve the underlying decision making process. Value of Information Analysis deals with the meta decision problem 2.

The underlying decision problem and its notational framework We are considering a decision maker who can influence an ecological-economic system having two alternative decisions d_1 and d_2 at hand. We assume, that the system can be characterized by the n -dimensional vector X . The characteristics X , are not necessarily known exactly to the decision maker. However, we assume furthermore that she is able to quantify this uncertainty which we call an *estimate* of the characteristics. Mathematically, an estimate is a random variable with probability density ρ_X .

Furthermore, the characteristics X determine the welfare W_d according to the welfare function w_d :

$$W_d = w_d(X)$$

Thus, the welfare of decision d is also a random variable which probability distribution we call ρ_{W_d} . The welfare function w_d values the decision d given a certain state X of the system. In other words, decision d_2 is preferred over decision d_1 , if and only if, the expected welfare of decision d_2 is greater than the expected welfare¹ of decision d_1 , formally

$$d_1 \prec d_2 \Leftrightarrow E[W_{d_1}] < E[W_{d_2}].$$

This means the best decision d^* is the one which maximizes welfare:

$$d^* := \arg \max_{d=d_1, d_2} E[W_d] \quad (\text{ES } 1)$$

This maximization principle has a dual minimization principle. We define the net benefit $\text{NB}_{d_1} := W_{d_1} - W_{d_2}$ as the difference between the welfare of the two decision alternatives. A loss L_d is characterized if a decision d produces a negative net benefit. No loss occurs if the decision produces a positive net benefit. This is reflected in the formal definition

$$L_d := \begin{cases} -\text{NB}_d & , \text{ if } \text{NB}_d < 0 \\ 0 & , \text{ otherwise.} \end{cases}$$

¹For a comprehensive discussion of the concept of social preference ordering and its representation by a welfare function cf. Gravelle and Rees (2004).

Using this notion it can be shown that the maximization of expected welfare is equivalent to the minimization of the expected loss $EL_d := E[L_d]$.

The meta problem of reducing uncertainty using the value of information The meta problem is how resources shall be allocated to reduce the uncertainty inherent to the underlying decision problem (eq. ES 1)., The uncertainty of the underlying decision problem is represented by the probability density ρ_X . Here, we interpret the probability density ρ_X of the characteristics X as the *information that is available of the ecological-economic system* that is to be manipulated by the decision maker.

Initially, the decision maker forms an *estimate* to quantify the current information $\rho_X^{current}$. She is interested on how to increase the current information by a particular investigation on the system, which we call a *measurement*, to improve the underlying decision making process. To this end, she specifies a prospective hypothetical information $\rho_X^{prospective}$ and assigns a value to this increase in information. The meta problem is solved by comparing this value of information to the prospective costs of the corresponding measurement.

Above, it was stated that the optimal decision d^* is the same if expected welfare is maximized or expected loss is minimized. However, the value of information analysis depends on if one focuses on risk management or on welfare maximization. Thus, we consider two different, but related, principles of defining the value of information: reducing minimal expected loss and increasing maximal expected welfare, which will be described in the sequel.

Meta decision principle: Focus on risk management or minimization of expected opportunity loss This principle is along the line described in Hubbard (2014). The Expected Opportunity Loss (EOL) is defined as the expected loss for the best decision. The best decision minimizes the expected loss:

$$EOL := \min \{EL_{d_1}, EL_{d_2}\}$$

The EOL is always conditional on the available information which is characterized by the probability distribution of X ρ_X : $EOL = EOL(\rho_X)$. The Expected Value of Information is the decrease in the EOL for an information improvement from the current ($\rho_X^{current}$) to a better prospective (hypothetical) information ($\rho_X^{prospective}$):

$$EVI := EOL(\rho_X^{current}) - EOL(\rho_X^{prospective}).$$

If some variables under $\rho_X^{prospective}$ are assumed to be known with certainty the EVI is called the *Clustered Expected Value of Perfect Information* (Clustered EVPI). If only one variable is assumed to be known with certainty, it is called the Individual EVPI. More precisely, if one assumes under $\rho_X^{prospective}$ to perfectly know (X_1, \dots, X_k) to equal (a_1, \dots, a_k) then one can specify the notation as

$$\text{Clustered EVPI}_{a_1, \dots, a_k}^{1, \dots, k} := EOL_{\text{current}} - EOL_{X_1=a_1, \dots, X_k=a_k}.$$

The superscripts of Clustered EVPI denote the components of the characteristics $X = (X_1, \dots, X_n)$ that are known with certainty. The subscripts denote the corresponding certain values. $EOL_{X_1=a_1, \dots, X_k=a_k}$ is the prospective expected opportunity loss with perfectly known characteristics X_1, \dots, X_k and the remaining characteristics X_{k+1}, \dots, X_n being distributed according to the current information $\rho_X^{current}$. Analogously is the Individual EVPI defined for knowing only one characteristic, here characteristic i exactly with value a_i :

$$\text{Individual EVPI}_{a_i}^i := EOL_{\text{current}} - EOL_{X_i=a_i}.$$

Thus, the EV(P)I depends on the model w_p for valuing a decision d , the current information $\rho_X^{current}$, i.e. the current estimate, and the specification of a prospective improvement in information $\rho_X^{prospective}$, i.e. a better estimate.

Meta decision principle: Focus on utility maximization or maximization of expected welfare Instead of founding the decision on loss minimization, one can found it on the value maximization (cf. Wikipedia: en.wikipedia.org/wiki/Expected_value_of_perfect_information). The Expected Value is the net benefit of a decision. The Expected Maximum Value (EMV) is the expected value for the best decision, where best, here, refers to the decision that maximizes the Expected Value. Analogously to the EOL case, the EMV is conditional on the available information (ρ_X): $EMV = EMV(\rho_X)$. In this case the Expected Value of Information is the increase in EMV for an information improvement from the status quo ($\rho_X^{current}$) to a better prospective (hypothetical) information ($\rho_X^{prospective}$): $EVI := EMV(\rho_X^{prospective}) - EMV(\rho_X^{current})$. Individual and Clustered EVPI are defined as in the other case. NB: Here, both definitions of the EVI deviate from another common definition, which takes the expectation of the Clustered EVPI over the distribution of the variables assumed to be known with certainty (cf. Jeffrey and Pannell (2013)). However, in practice computing this further expectation value is computationally very expensive. Thus, we will use the procedure described above as an approximation. (-#LG: check this last paragraph)

Solving the practical problem of calculating expected welfare by Monte Carlo simulation In order to solve the underlying decision problem the expected welfare $E[W_d]$ or expected loss EL_d has to be calculated for each decision $d = d_1, d_2$. *A priori*, only the probability density of X , viz. ρ_X , is known. Although this determines expected welfare by

$$E[W_d] = E_{\rho_X}[w_d] = \int w_d(x) \rho_X(x) dx,$$

the integral might be difficult (or impossible) to solve directly in practical applications. If one would have a random sample w_1, \dots, w_N of the welfare W_d , the expected welfare could be approximated as the mean of the random sample:

$$E[W_d] \approx \frac{1}{N} \sum_{i=1}^N w_i.$$

The Monte Carlo simulation is the method to generate the random sample of W_d as follows. First, we sample N values of X : X_1, \dots, X_N . Then we calculate the corresponding welfare $w_1 := w_d(X_1), \dots, w_N := w_d(X_N)$. Applying the law of large numbers the sample mean $\bar{w}(N)$ converges to the expectation value

$$\bar{w}(N) := \frac{1}{N} \sum_{i=1}^N w_i \xrightarrow{N \rightarrow \infty} \int w \rho_{W_d}(w) dw.$$

Let's call the probability density of the simulated sample $\rho_W^{MC}(N)$. According to the law of large numbers, for large N the probability density obtained by the Monte Carlo simulation converges to the exact probability density:

$$\rho_W^{MC}(N) \xrightarrow{N \rightarrow \infty} \rho_{W_d}.$$

Contents

1	Introduction	4
2	The Underlying Decision Making Principle of a Binary Decision Problem	4
2.1	Special case: Status quo and project approval	8

3	The Meta Decision of Reducing Uncertainty Using the Expected Value of Information	9
3.1	Meta decision principle: Focus on Risk Management or Minimization of Expected Opportunity Loss	10
3.1.1	Expected Value of Perfect Information	10
3.1.2	Special case: Status quo and project approval	12
3.2	Meta decision principle: Focus on utility maximization or maximization of expected welfare	14
4	Solving the Practical Problem of Calculating Expectation Values by Monte Carlo Simulation	14
4.1	The Practical Problem of Calculating Expectation Values	14
4.2	Monte Carlo Simulations	15
4.3	Solving the Practical Problem of Calculating Expectation Values by Monte Carlo Simulation	16
5	Conclusion	16

1 Introduction

The topic of this analysis is the mathematical formalism of the Applied Information Economics (AIE) approach to decision making developed in Hubbard (2014). An important part of the AIE is a quantitative analysis of welfare based decision making processes² using Monte Carlo simulations. This is the focus of this article. These decision making processes can be categorized into two levels of decision making:

1. The actual problem of interest of a policy maker which we call the *underlying welfare based decision* on how to influence an ecological-economic system based on a particular information on the system available to the decision maker and
2. the *meta decision* on how to allocate resources to reduce the uncertainty in the underlying decision problem, i.e to increase the current information to improve the underlying decision making process.

The first problem, i.e. the underlying problem, is the problem of choosing the decision which maximizes expected welfare. The welfare function can be interpreted as a von Neumann-Morgenstern utility function (Gravelle and Rees, 2004). Whereas, the second problem, i.e. the meta decision problem, is dealt with using the *Value of Information Analysis (VIA)*. Value of Information Analysis seeks to assign a value to a certain reduction in uncertainty or, equivalently, increase in information. Uncertainty is dealt with in a probabilistic manner. Probabilities are transformed via Monte Carlo simulations.

2 The Underlying Decision Making Principle of a Binary Decision Problem

In this discussion, we are considering a decision maker who can influence an ecological-economic system having two alternative decisions (or projects) d_1 and d_2 at hand. The set of these two decisions is called

$$\mathcal{D} = \{d_1, d_2\}. \quad (1)$$

²cf. footnote 1

Assume there is a welfare function³

$$\begin{aligned} w_d : \Xi \subset \mathbb{R}^n &\longrightarrow \mathbb{R} \\ x &\mapsto w_d(x) \end{aligned} \quad (2)$$

which orders the decisions for the system with characteristics x according to

$$d_1 \prec d_2 \Leftrightarrow w_{d_1}(x) < w_{d_2}(x). \quad (3)$$

This means decision d_2 is strictly preferred over its alternative d_1 if, and only if, the welfare under decision d_2 is strictly greater than the welfare under decision d_1 .

The characteristics X , are not necessarily known exactly to the decision maker. However, we assume furthermore that she is able to quantify this uncertainty which we call an *estimate* of the characteristics.

Definition 2.1 (Estimate). The knowledge about the ecological-economic system which is available to the decision maker is called an *estimate* and is formally described by the probability space (Ω, \mathcal{A}, P) where $X : \Omega \longrightarrow \mathbb{R}^n$ with $\Xi = X(\Omega)$ is the random variable of realized characteristics of the ecological-economic system and $P_X = X(P)$ the probability distribution of X . A particular characteristic is realized as $x = X(\omega)$.

Now define $W_d := w_d \circ X$. Thus, the welfare of decision d is a random variable with distribution $w_d(P_X)$:

$$W_d \sim w_d(P_X) := P_{W_d} \quad (4)$$

This completes the notation to define the decision making principle.

Definition 2.2 (The Decision Making Principle (Maximization of Expected Welfare)). Decision d_2 is strictly preferred over its alternative d_1 if, and only if, the expected welfare under decision d_2 is strictly greater than the expected welfare under decision d_1 :

$$d_1 \prec d_2 \Leftrightarrow \mathbb{E}[W_{d_1}] < \mathbb{E}[W_{d_2}], \quad (5)$$

where

$$\mathbb{E}[W_d] = \int_{\Omega} W_d dP \quad (6)$$

is the expected welfare of decision $d \in \mathcal{D}$. This implies that the optimal decision d^* maximizes expected welfare:

$$d^* = \arg \max_{d \in \mathcal{D}} \mathbb{E}[W_d] \quad (7)$$

Note.

$$\mathbb{E}[W_d] = \int_{\Omega} W_d dP = \int_{\Omega} w_d \circ X dP = \int_{X(\Omega)} w_d dP_X = \int w_d(x) P_X(dx) = \mathbb{E}_{P_X}[w_d] \quad (8)$$

Below we will show that maximization of expected welfare is equivalent to minimization of expected loss and maximization of expected net benefit. The notions of loss and net benefit are specified in the sequel.

³cf. footnote 1

Definition 2.3 (Loss function and Loss). The *loss function* for decision d :

$$l_d := \begin{cases} -(w_d - w_{d'}) & , \text{ if } w_d - w_{d'} < 0 \\ 0 & , \text{ otherwise} \end{cases}, d \neq d' \quad (9)$$

$$= -\chi_{\{w_d - w_{d'} < 0\}} (w_d - w_{d'}), \quad d \neq d' \quad (10)$$

where χ is the characteristic function

$$\chi_A(x) := \begin{cases} 1 & , \text{ if } x \in A \\ 0 & , \text{ otherwise} \end{cases}. \quad (11)$$

The *loss* for decision d is the random variable

$$\begin{aligned} L_d &:= l_d \circ X \\ &= \begin{cases} -(W_d - W_{d'}) & , \text{ if } W_d - W_{d'} < 0 \\ 0 & , \text{ otherwise} \end{cases}, d \neq d'. \end{aligned} \quad (12)$$

Note. The loss is defined being non-negative, i.e.

$$L_d \geq 0. \quad (13)$$

Definition 2.4 (Expected Loss). *Expected loss* is

$$\text{EL}_d := \text{E}[L_d]. \quad (14)$$

Note.

$$\text{EL}_d \geq 0. \quad (15)$$

Proposition 2.5.

$$\text{EL}_d = \text{E}_{P_X}[l_d] \quad (16)$$

Proof. It holds for $d \neq d'$ that

$$\begin{aligned} \text{EL}_d &= \int L_d dP \\ &= - \int_{\{W_d - W_{d'} < 0\}} (W_d - W_{d'}) dP \\ &= - \int_{\{(w_d - w_{d'}) \circ X < 0\}} (w_d - w_{d'}) \circ X dP \\ &= - \int_{\{w_d - w_{d'} < 0\}} (w_d - w_{d'}) d(X(P)) \\ &= - \int_{\{w_d - w_{d'} < 0\}} (w_d - w_{d'}) dP_X \\ &= - \int_{\{w_d - w_{d'} < 0\}} (w_d(x) - w_{d'}(x)) P_X(dx) \\ &= \int l_d(x) P_X(dx) \\ &= \text{E}_{P_X}[l_d] \end{aligned} \quad \square$$

Definition 2.6 (Net Benefit). The *net benefit* of decision d vs. the alternative d' is

$$\text{NB}_d := W_d - W_{d'}, \quad d' \neq d. \quad (17)$$

Note. Now the loss can be interpreted as the absolute value of the net benefit where it is negative. Positive net benefit means zero loss:

$$L_d = \begin{cases} -\text{NB}_d & , \text{ if } \text{NB}_d < 0 \\ 0 & , \text{ otherwise} \end{cases} \quad (18)$$

In other words, loss occurs only in case of negative net benefit.

Definition 2.7 (Expected Net Benefit). The *expected net benefit* of decision d vs. the alternative d' is

$$\text{ENB}_d := \text{E} [\text{NB}_d] = \text{E} [W_d - W_{d'}], \quad d' \neq d. \quad (19)$$

Note. It holds for $d \neq d'$ that

$$\text{ENB}_d = -\text{ENB}_{d'}. \quad (20)$$

Proposition 2.8. It holds for $d \neq d'$ that

$$\text{EL}_d + \text{ENB}_d = \text{EL}_{d'}. \quad (21)$$

Proof.

$$\begin{aligned} \text{EL}_d &= \int L_d dP \\ &= \int_{\{W_d - W_{d'} < 0\}} (W_{d'} - W_d) dP \\ &= \int_{\{W_d - W_{d'} < 0\}} W_{d'} dP - \int_{\{W_d - W_{d'} < 0\}} W_d dP \\ &= \int W_{d'} dP - \int_{\{W_d - W_{d'} \geq 0\}} W_{d'} dP + \int_{\{W_d - W_{d'} \geq 0\}} W_d dP - \int W_d dP \\ &= - \int (W_d - W_{d'}) dP - \int_{\{W_d - W_{d'} \geq 0\}} (W_{d'} - W_d) dP \\ &= - \underbrace{\int (W_d - W_{d'}) dP}_{\text{ENB}_d} - \underbrace{\int_{\{W_{d'} - W_d < 0\}} (W_{d'} - W_d) dP}_{-\text{EL}_{d'}} \quad \square \end{aligned}$$

Theorem 2.9 (Equivalent Formulations of the Decision Making Principle). *The statement that decision d' is strictly better than the alternative d , i.e $d \prec d'$, is equivalent to each of the following*

- (i) $\text{E} [W_d] < \text{E} [W_{d'}]$
- (ii) $\text{ENB}_{d'} > 0$
- (iii) $\text{EL}_d > \text{EL}_{d'}$.

Proof. Ad (i): This is the definition of the decision making principle (eq. 5).

Ad (ii): $\text{E} [W_d] < \text{E} [W_{d'}] \iff \text{ENB}_{d'} = \text{E} [W_{d'}] - \text{E} [W_d] > 0$.

Ad (iii): Due to prop. 2.8 (ii) is equivalent to $\text{ENB}_{d'} = \text{EL}_d - \text{EL}_{d'} > 0$. \square

Corollary 2.10. Expected Net Benefit Maximization:

$$d^* = \arg \max_{d \in \mathcal{D}} \text{ENB}_d = \begin{cases} d_1 & \text{if } \text{ENB}_{d_1} > 0 \\ d_2 & \text{otherwise} \end{cases} \quad (22a)$$

Expected Loss Minimization:

$$d^* = \arg \min_{d \in \mathcal{D}} \text{EL}_d = \begin{cases} d_1 & \text{if } \text{EL}_{d_1} < \text{EL}_{d_2} \\ d_2 & \text{otherwise} \end{cases} \quad (22b)$$

2.1 Special case: Status quo and project approval

A very common actual binary decision problem is the question if a certain project shall be approved versus continuing with business as usual, i.e. the status quo. It appears to be natural to identify the status quo with zero welfare. This is a special case⁴ of the binary decision problem (eq. 1) that we are considering here. The two decision alternatives are d_1 : project approval (PA) and d_2 : status quo (SQ), formally:

$$\mathcal{D} = \{\text{PA}, \text{SQ}\} \quad \text{and} \quad (23)$$

the welfare of the approved project (or project outcome or yield of the project) is the random variable W_{PA} with distribution $P_{W_{\text{PA}}} = w_{\text{PA}}(P_X)$:

$$W_{\text{PA}} \sim w_{\text{PA}}(P_X) = P_{W_{\text{PA}}} \quad (24)$$

and the welfare of the status quo serves as reference and is normalized to zero:

$$W_{\text{SQ}} \equiv 0, \quad (25)$$

which implies zero expected welfare of the status quo:

$$\text{E}[W_{\text{SQ}}] = 0. \quad (26)$$

The loss functions simply are

$$L_{\text{PA}} = \begin{cases} -W_{\text{PA}} & , \text{ if } W_{\text{PA}} < 0 \\ 0 & , \text{ otherwise} \end{cases} \quad (27a)$$

and

$$L_{\text{SQ}} = \begin{cases} W_{\text{PA}} & , \text{ if } W_{\text{PA}} > 0 \\ 0 & , \text{ otherwise.} \end{cases} \quad (27b)$$

The normalization of the status quo's welfare to zero implies that the net benefit of the project approval simply equals its benefit and that the net benefit of the status quo equals the negative welfare of the approved project (follows directly from defn. 2.6):

$$\text{NB}_{\text{PA}} = W_{\text{PA}} \quad (28a)$$

$$\text{NB}_{\text{SQ}} = -W_{\text{PA}} \quad (28b)$$

⁴Actually, one can show, that this special case is equivalent to the discussion above.

Prop. 2.8 reads

$$EL_{PA} + ENB_{PA} = EL_{SQ} \quad (29a)$$

$$EL_{SQ} + ENB_{SQ} = EL_{PA} \quad (29b)$$

or due to $ENB_{SQ} = -E[W_{PA}] = -ENB_{PA}$ simply

$$EL_{PA} + E[W_{PA}] = EL_{SQ}. \quad (30)$$

Finally, the decision making principle reads in these terms, that the project is approved, if and only if, its expected welfare is positive:

$$PA \succ SQ \Leftrightarrow E[W_{PA}] > 0 \quad (31)$$

Then, the optimal decision is

$$d^* = \begin{cases} PA & \text{if } E[W_{PA}] > 0 \\ SQ & \text{otherwise} \end{cases} \quad (32a)$$

or equivalently

$$d^* = \begin{cases} PA & \text{if } EL_{PA} < EL_{SQ} \\ SQ & \text{otherwise.} \end{cases} \quad (32b)$$

Summarizing, choosing positive welfare is equivalent to minimizing expected loss.

Example 2.11 (Net Present Value). A net present value (NPV) analysis is the most important practical example of the status quo and project approval case (eq. 25). The definition of the net present value of the project approval:

$$NPV := \sum_{t=0}^{\infty} \frac{b_t - c_t}{(1 + \delta)^{t-1}} \quad (33)$$

is the discounted sum of the difference between all benefits b_t and costs c_t in period t with discount rate δ . This implies

$$W_{PA} = NPV \quad (34a)$$

$$W_{SQ} = 0 \quad (34b)$$

and thus fulfills the defining equation (eq. 25).

3 The Meta Decision of Reducing Uncertainty Using the Expected Value of Information

The meta problem is how resources shall be allocated to reduce the uncertainty inherent to the underlying decision problem (eq. 1)., The uncertainty of the underlying decision problem is represented by the probability distribution P_X . Here, we interpret the probability distribution P_X of the characteristics X as the *information that is available of the ecological-economic system* that is to be manipulated by the decision maker.

Initially, the decision maker forms an *estimate* to quantify the current information $P_X^{current}$. She is interested on how to increase the current information by a particular investigation on the

system, which we call a *measurement*, to improve the underlying decision making process. To this end, she specifies a prospective hypothetical information $P_X^{\text{prospective}}$ and assigns a value to this increase in information. The meta problem is solved by comparing this value of information to the prospective costs of the corresponding measurement.

Above, it was stated that the optimal decision d^* is the same if expected welfare is maximized or expected loss is minimized. However, the value of information analysis depends on if one focuses on risk management or on welfare maximization. Thus, we consider two different, but related, principles of defining the value of information: reducing minimal expected loss and increasing maximal expected welfare, which will be described in the sequel.

3.1 Meta decision principle: Focus on Risk Management or Minimization of Expected Opportunity Loss

Cor. 2.10 motivates the following

Definition 3.1 (Expected Opportunity Loss (EOL)). The *expected opportunity loss (EOL)* of a given information P is the expected loss of the best decision given this information:

$$\text{EOL}(P) := \min_{d \in \mathcal{D}} \text{EL}_d \quad (35)$$

Proposition 3.2.

$$\text{EOL}(P) = \min \{ \text{EL}_{d_1}, \text{ENB}_{d_1} + \text{EL}_{d_1} \} = \min \{ \text{EL}_{d_2}, \text{ENB}_{d_2} + \text{EL}_{d_2} \} \quad (36)$$

Proof. Follows directly from the definition of the expected opportunity loss (defn. 3.1) and prop. 2.8. \square

Definition 3.3 (Expected Value of Information). Let the current information on the characteristics of the system be described by the probability distribution P^{current} and the prospective information by a certain investigation on the system which induces the change in information ΔP of the system be $P^{\text{prospective}}$. Let the corresponding opportunity losses be denoted by

$$\text{EOL}_{\text{current}} = \min_{d \in \mathcal{D}} \int L_d dP^{\text{current}} \quad (37a)$$

and

$$\text{EOL}_{\text{prospective}} = \min_{d \in \mathcal{D}} \int L_d dP^{\text{prospective}}, \quad (37b)$$

respectively. Then, the *Expected Value of Information (EVI)* of this information change ΔP is defined as

$$\text{EVI}(\Delta P) := \text{EOL}_{\text{current}} - \text{EOL}_{\text{prospective}}. \quad (38)$$

3.1.1 Expected Value of Perfect Information

This subsection deals with the cases that the prospective information $P^{\text{prospective}}$ implies certainty of several or all components of the characteristics x of the system under study. Formally,

$$P_X^{\text{prospective}} = P_X^{[x_1=\bar{x}_1, \dots, x_n=\bar{x}_n]} := \delta_{\bar{x}_1} \otimes \dots \otimes \delta_{\bar{x}_k} \otimes P_{(X_j)_{j=k+1, \dots, n}}^{\text{current}}, \quad (39)$$

where, w.l.o.g.⁵ the first k components of the characteristics x are known with certainty, viz. $\bar{x}_1, \dots, \bar{x}_k$, $\delta_{\bar{x}_i}$ is the Dirac-distribution⁶ and $P_{(X_j)_{j=k+1, \dots, n}}^{\text{current}}$ is the current probability distribution of the remaining components of x , i.e. x_{k+1}, \dots, x_n . In other words, $P_X^{[x_1=\bar{x}_1, \dots, x_n=\bar{x}_n]}$ means certainty about the first k characteristics and the same uncertainty as under the current information for the remaining characteristics x_{k+1}, \dots, x_n .

Definition 3.4 (Expected Value of Perfect Information (EVPI)). The EVI is called *Expected Value of Perfect Information (EVPI)* if all characteristics x_1, \dots, x_n are assumed to be known with their exact values $\bar{x}_1, \dots, \bar{x}_n$, respectively:

$$\text{EVPI}[\bar{x}_1, \dots, \bar{x}_n] := \text{EOL}_{\text{current}} - \text{EOL}_{[x_1=\bar{x}_1, \dots, x_n=\bar{x}_n]}, \quad (40)$$

where

$$\text{EOL}_{[x_1=\bar{x}_1, \dots, x_n=\bar{x}_n]} := \min_{d \in \mathcal{D}} \mathbb{E}_{x_1=\bar{x}_1, \dots, x_n=\bar{x}_n} [L_d] \quad (41)$$

and

$$\mathbb{E}_{x_1=\bar{x}_1, \dots, x_n=\bar{x}_n} [L_d] := \int l_d d\delta_{\bar{x}_1} \otimes \dots \otimes \delta_{\bar{x}_n} = l_d(\bar{x}_1, \dots, \bar{x}_n). \quad (42)$$

Definition 3.5 (Clustered EVPI first k characteristics). The EVI is called *Clustered EVPI* of characteristics $1, \dots, k$ at the respective values $\bar{x}_1, \dots, \bar{x}_k$ if k characteristics x_1, \dots, x_k are assumed to be known with their exact values $\bar{x}_1, \dots, \bar{x}_k$, respectively:

$$\text{Clustered EVPI}[\bar{x}_1, \dots, \bar{x}_k] := \text{EOL}_{\text{current}} - \text{EOL}_{x_1=\bar{x}_1, \dots, x_k=\bar{x}_k}, \quad (43)$$

where

$$\text{EOL}_{x_1=\bar{x}_1, \dots, x_k=\bar{x}_k} := \min_{d \in \mathcal{D}} \mathbb{E}_{x_1=\bar{x}_1, \dots, x_k=\bar{x}_k} [L_d] \quad (44)$$

and

$$\mathbb{E}_{x_1=\bar{x}_1, \dots, x_k=\bar{x}_k} [L_d] := \int l_d d\delta_{\bar{x}_1} \otimes \dots \otimes \delta_{\bar{x}_k} \otimes P_{(X_j)_{j=k+1, \dots, n}}^{\text{current}} = \int l_d(\bar{x}_1, \dots, \bar{x}_k, x) P_{(X_j)_{j=k+1, \dots, n}}^{\text{current}}(dx). \quad (45)$$

Definition 3.6 (Individual EVPI for the first characteristic). The Clustered EVPI is called *Individual EVPI* of characteristic 1 at the value \bar{x}_1 if characteristic x_1 is assumed to be known with the exact value \bar{x}_1 :

$$\text{Individual EVPI}[\bar{x}_1] := \text{EOL}_{\text{current}} - \text{EOL}_{x_1=\bar{x}_1}, \quad (46)$$

where

$$\text{EOL}_{x_1=\bar{x}_1} := \min_{d \in \mathcal{D}} \mathbb{E}_{x_1=\bar{x}_1} [L_d] \quad (47)$$

and

$$\mathbb{E}_{x_1=\bar{x}_1} [L_d] := \int l_d d\delta_{\bar{x}_1} \otimes P_{(X_j)_{j=2, \dots, n}}^{\text{current}} = \int l_d(\bar{x}_1, x) P_{(X_j)_{j=2, \dots, n}}^{\text{current}}(dx). \quad (48)$$

⁵without loss of generality

⁶The Dirac distribution is defined as

$$\delta_a(A) := \begin{cases} 1 & a \in A \\ 0 & \text{otherwise.} \end{cases}$$

It means that the event a will occur with certainty.

Definition 3.7 (Clustered EVPI k arbitrary characteristics).

$$\text{Clustered EVPI}_{\bar{x}_{i_1}, \dots, \bar{x}_{i_k}}^{i_1, \dots, i_k} := \text{EOL}_{\text{current}} - \text{EOL}_{x_{i_1}=\bar{x}_{i_1}, \dots, x_{i_k}=\bar{x}_{i_k}}, \quad (49)$$

where

$$\text{EOL}_{x_{i_1}=\bar{x}_{i_1}, \dots, x_{i_k}=\bar{x}_{i_k}} := \min_{d \in \mathcal{D}} \mathbb{E}_{x_{i_1}=\bar{x}_{i_1}, \dots, x_{i_k}=\bar{x}_{i_k}} [L_d] \quad (50)$$

and

$$\mathbb{E}_{x_{i_1}=\bar{x}_{i_1}, \dots, x_{i_k}=\bar{x}_{i_k}} [L_d] := \int l_d(x_1, \dots, x_{i_1-1}, \bar{x}_{i_1}, x_{i_1+1}, \dots, x_{i_k-1}, \bar{x}_{i_k}, x_{i_k+1}, \dots, x_n) \times P_{(X_j)_{j=1, \dots, i_1-1, i_1+1, \dots, i_k-1, i_k+1, \dots, n}}^{\text{current}}(\mathrm{d}\mathbf{x}) \quad (51)$$

Definition 3.8 (Individual EVPI for an arbitrary characteristic).

$$\text{Individual EVPI}_{\bar{x}_i}^i := \text{EOL}_{\text{current}} - \text{EOL}_{x_i=\bar{x}_i}, \quad (52)$$

where

$$\text{EOL}_{x_i=\bar{x}_i} := \min_{d \in \mathcal{D}} \mathbb{E}_{x_i=\bar{x}_i} [L_d] \quad (53)$$

and

$$\mathbb{E}_{x_i=\bar{x}_i} [L_d] := \int l_d(x_1, \dots, x_{i-1}, \bar{x}_i, x_{i+1}, \dots, x_n) P_{(X_j)_{j=1, \dots, i-1, i+1, \dots, n}}^{\text{current}}(\mathrm{d}\mathbf{x}) \quad (54)$$

Summarizing, if some variables under $P^{\text{prospective}}$ are assumed to be known with certainty the EVI is called the *Clustered Expected Value of Perfect Information* (Clusterd EVPI). If only one variable is assumed to be known with certainty, it is called the Individual EVPI. More precisely, if one assumes under $P^{\text{prospective}}$ to perfectly know $(x_{i_1}, \dots, x_{i_k})$ to equal $(\bar{x}_{i_1}, \dots, \bar{x}_{i_k})$ then one can specify the notation as Clustered EVPI $[\bar{x}_{i_1}, \dots, \bar{x}_{i_k}]$. Analogously, Individual EVPI $[x_i = a_i]$. Thus, the EV(P)I depends on the model w_d for valuing decision d , the status quo information, i.e. the current information or estimate P^{current} , and the specification of a hypothetical improvement in information, i.e. a better estimate which we call $P^{\text{prospective}}$.

3.1.2 Special case: Status quo and project approval

$$\text{EOL} = \min \{ \text{EL}_{\text{PA}}, \text{EL}_{\text{PA}} + \mathbb{E}[W_{\text{PA}}] \} \quad (55)$$

Proposition 3.9. For $\mathcal{D} = \{\text{PA}, \text{SQ}\}$ and $w_{\text{SQ}} \equiv 0$ it holds that

$$\text{EOL}_{[x_1=\bar{x}_1, \dots, x_n=\bar{x}_n]} = 0. \quad (56)$$

Proof. Recall

$$l_{\text{PA}} = \begin{cases} -w_{\text{PA}} & , \text{ if } w_{\text{PA}} < 0 \\ 0 & , \text{ otherwise} \end{cases}$$

and

$$l_{\text{SQ}} = \begin{cases} w_{\text{PA}} & , \text{ if } w_{\text{PA}} > 0 \\ 0 & , \text{ otherwise.} \end{cases}$$

It follows

$$\begin{aligned} \mathbb{E}_{x_1=\bar{x}_1, \dots, x_n=\bar{x}_n} [L_{\text{PA}}] &= \begin{cases} -w_{\text{PA}}(\bar{x}_1, \dots, \bar{x}_n) & , \text{ if } w_{\text{PA}}(\bar{x}_1, \dots, \bar{x}_n) < 0 \\ 0 & , \text{ otherwise} \end{cases} \quad \text{and} \\ \mathbb{E}_{x_1=\bar{x}_1, \dots, x_n=\bar{x}_n} [L_{\text{SQ}}] &= \begin{cases} w_{\text{PA}}(\bar{x}_1, \dots, \bar{x}_n) & , \text{ if } w_{\text{PA}}(\bar{x}_1, \dots, \bar{x}_n) > 0 \\ 0 & , \text{ otherwise.} \end{cases} \end{aligned}$$

And finally

$$\text{EOL}_{[x_1=\bar{x}_1, \dots, x_n=\bar{x}_n]} = \min_{d \in \{\text{SQ}, \text{PA}\}} l_d(\bar{x}_1, \dots, \bar{x}_n) = 0.$$

□

Proposition 3.9 directly leads to the following

Corollary 3.10. *For $\mathcal{D} = \{\text{PA}, \text{SQ}\}$ and $w_{\text{SQ}} \equiv 0$ it holds that*

$$\text{EVPI}[\bar{x}_1, \dots, \bar{x}_n] = \text{EOL}_{\text{current}}. \quad (57)$$

The defining normalization (eq. 25), which results in eq. (26), has the important consequence that the expected value of perfect information equals the expected opportunity loss under the current information P_{current} .

Definition 3.11 (Approximate EVPI). For $\mathcal{D} = \{\text{PA}, \text{SQ}\}$, $w_{\text{SQ}} \equiv 0$ and $n \geq 2$ we define

$$\text{Approximate EVPI}^i := \text{EOL}^i - \text{EOL}_{[x_1=\mathbb{E}[x_1], \dots, x_n=\mathbb{E}[x_n]]} = \text{EOL}^i, \quad (58)$$

where

$$\text{EOL}^i := \text{EOL}_{[x_1=\mathbb{E}[x_1], \dots, x_{i-1}=\mathbb{E}[x_{i-1}], x_{i+1}=\mathbb{E}[x_{i+1}], \dots, x_n=\mathbb{E}[x_n]]}. \quad (59)$$

Example 3.12 (Bernoulli uncertainty (cf. Hubbard (2014, pp.147, in particular exhibit 7.1))). Consider a decision between investing in a new advertising campaign (PA) or not investing in it (SQ). Furthermore, we restrict the model to two possible events: Either the campaign works ($x = 1$) with probability p_1 or the campaign does not work ($x = 0$) with probability p_0 . If the project is approved (PA) and the campaign works it produces a gain $w_{\text{PA}}(1) = w_{\text{PA}}^1 > 0$, if it fails, you lose your investment $w_{\text{PA}}(0) = w_{\text{PA}}^0 < 0$. If the project is not approved (SQ) in neither case you neither lose nor gain. On the one hand, if your campaign would not have worked you would not have invested and thus you would not have incurred a loss: $w_{\text{SQ}}(0) = 0$. On the other hand, if your campaign would have worked you could not harvest the gain because you rejected the project: $w_{\text{SQ}}(1) = 0$. Summarizing:

$$\Xi := \{0, 1\} \quad (60a)$$

$$P_X^{\text{current}}(\{0\}) =: p_0 \quad (60b)$$

$$P_X^{\text{current}}(\{1\}) =: p_1 \quad (60c)$$

$$w_{\text{PA}}(0) = w_{\text{PA}}^0 < 0 \quad (60d)$$

$$w_{\text{PA}}(1) = w_{\text{PA}}^1 > 0 \quad (60e)$$

$$w_{\text{SQ}}(0) = 0 \quad (60f)$$

$$w_{\text{SQ}}(1) = 0 \quad (60g)$$

From this the loss function can be calculated as

$$l_{PA}(0) = -(w_{PA}^0 - 0) = -w_{PA}^0 > 0 \quad (61a)$$

$$l_{PA}(1) = 0 \quad (61b)$$

$$l_{SQ}(0) = 0 \quad (61c)$$

$$l_{SQ}(1) = -(0 - w_{PA}^1) = w_{PA}^1 > 0 \quad (61d)$$

and the expected losses for the two alternatives

$$E[L_{PA}] = p_0 l_{PA}(0) + p_1 l_{PA}(1) = -p_0 w_{PA}^0 \quad (62a)$$

$$E[L_{SQ}] = p_0 l_{SQ}(0) + p_1 l_{SQ}(1) = p_1 w_{PA}^1. \quad (62b)$$

Finally, the expected opportunity loss at the current information is

$$EOL_{current} = \min \{ -p_0 w_{PA}^0, p_1 w_{PA}^1 \}, \quad (63)$$

which equals the EVPI (cor. 3.10):

$$EVPI[\bar{x}] = \min \{ -p_0 w_{PA}^0, p_1 w_{PA}^1 \}. \quad (64)$$

for $\bar{x} = 0, 1$.

3.2 Meta decision principle: Focus on utility maximization or maximization of expected welfare

Instead of founding the decision on loss minimization, one can found it on the value maximization (cf. Wikipedia: en.wikipedia.org/wiki/Expected_value_of_perfect_information). The Expected Value is the net benefit of a decision. The Expected Maximum Value (EMV) is the expected value for the best decision, where best, here, refers to the decision that maximizes the Expected Value. Analogously to the EOL case, the EMV is conditional on the available information (P_X): $EMV = EMV(P_X)$. In this case the Expected Value of Information is the increase in EMV for an information improvement from the status quo ($\rho_X^{current}$) to a better prospective (hypothetical) information ($P_X^{prospective}$): $EVI := EMV(P_X^{prospective}) - EMV(P_X^{current})$. Individual and Clustered EVPI are defined as in the other case. NB: Here, both definitions of the EVI deviate from another common definition, which takes the expectation of the Clustered EVPI over the distribution of the variables assumed to be known with certainty (cf. Jeffrey and Pannell (2013)). However, in practice computing this further expectation value is computationally very expensive. Thus, we will use the procedure described above as an approximation. (-#LG: check this last paragraph)

4 Solving the Practical Problem of Calculating Expectation Values by Monte Carlo Simulation

4.1 The Practical Problem of Calculating Expectation Values

The decision analysis based on the minimization of the EOL and the quantification of the EVI of a prospective improvement of information require the calculation of expectation values such as

the expected loss $EL_d = E[L_d]$ or the expected welfare $E[W_d]$ of a certain decision d . Consider the calculation of the expected loss

$$EL_d = E[L_d] = \int l_d(x) P_X(dx). \quad (65)$$

Assume that the probability distribution P_X of X can be described by the probability density ρ_X , which means that we have to calculate the integral

$$EL_d = \int l_d(x) \rho_X(x) dx. \quad (66)$$

In practice, the problem arises that, this is often a high dimensional integral that is analytically not or only hard to solve. Also numerically, it might be very expensive or impossible to solve. Monte Carlo simulations are a way to tackle this problem.

4.2 Monte Carlo Simulations

A Monte Carlo simulation is a method that solves the following problem. Given a multivariate random variable $X = (X_1, \dots, X_k)$ with joint probability distribution P , i.e.

$$X \sim P_X. \quad (67)$$

Then the continuous function

$$f : R^k \rightarrow R^l \quad (68)$$

defines another random variable

$$Y := f \circ X \quad (69)$$

with distribution

$$Y \sim f(P_X). \quad (70)$$

Given there is a probability density ρ_X of X that defines P_X ⁷:

$$P_X(A) = \int_A \rho_X(x) dx \quad (71)$$

Then, there exists also a probability density ρ_Y for $P_Y = f(P)$:

$$P_Y(B) = \int_B \rho_Y(y) dy \quad (72)$$

The problem is the determination of the probability density ρ_y from the knowledge of ρ_x and f . A Monte Carlo simulation solves this problem approximately. It samples values for X which are plugged into the transformation function f , which form a simulated sample of Y . From this sample of Y an approximation of the probability density ρ_Y can be easily obtained by counting.

Let's describe this in more detail. We draw N times a realization of X which we call x_1, \dots, x_N , i.e. we draw a random sample of size N of X . From this, we obtain the simulated random sample

$$y_1 := f(x_1), \dots, y_N := f(x_N). \quad (73)$$

⁷This requires P_X to be absolutely continuous.

with probability density $\rho_Y^{MC}(N)$. According to the law of large numbers, for large N the probability density obtained by the Monte Carlo simulation converges to the exact probability density:

$$\rho_Y^{MC}(N) \xrightarrow{N \rightarrow \infty} \rho_Y. \quad (74)$$

4.3 Solving the Practical Problem of Calculating Expectation Values by Monte Carlo Simulation

The integral for EL_d (eq. 65) can be transformed using the well known theorem for the transformation of integrals (e.g. cf. Bauer (1992)) to

$$EL_d = \int l \cdot l_d(P_X)(d\mathbf{l}) = \int l P_{l_d \circ X}(d\mathbf{l}) = \int l P_{L_d}(d\mathbf{l}) = \int l \rho_{L_d}(l) d\mathbf{l}. \quad (75)$$

This is only a one dimensional integral. However, so far, it requires the knowledge of the distribution of L_d . We circumvent the direct calculation of P_{L_d} by its approximation using a Monte Carlo simulation. First, we sample N values of X : x_1, \dots, x_N . Then we calculate the corresponding losses $l_1 := l_d(x_1), \dots, l_N := l_d(x_N)$. Applying the law of large numbers the sample mean $\bar{l}(N)$ converges to the expectation value:

$$\bar{l}(N) := \frac{1}{N} \sum_{i=1}^N l_i \xrightarrow{N \rightarrow \infty} \int l \rho_{L_d}(l) d\mathbf{l} \quad (76)$$

or

$$EL_d \approx \frac{1}{N} \sum_{i=1}^N l_i \quad (77)$$

5 Conclusion

References

- Heinz Bauer. *Maß- und Integrationstheorie*. de Gruyter, Berlin; New York, 2nd edition, 1992.
- Hugh Gravelle and Ray Rees. *Microeconomics*. Pearson Education Limited, 3rd edition, 2004.
- Douglas W. Hubbard. *How to Measure Anything? - Finding the Value of "Intangibles" in Business*. John Wiley & Sons, Hoboken, New Jersey, 3rd edition, 2014.
- Scott R. Jeffrey and David J Pannell. Economics of prioritising environmental research: An expected value of partial perfect information (evppi) framework. Working paper, School of Agricultural and Resource Economics, University of Western Australia, 2013. URL <http://pur1.umn.edu/144944>.