

Kimi Ylilammi

A simulation tool for project portfolio selection

School of Science

Independent research project in applied mathematics
(Mat-2.4108)

Espoo 21.1.2014

Assignment supervisor:

Prof. Ahti Salo

Assignment advisor:

Lic.Sc.(Tech) Eeva Vilkkumaa

Contents

Table of contents	ii
Symbols and abbreviations	iii
1 Introduction	1
2 Project funding model	1
2.1 Value model	2
2.2 Model for the single-period simulation	2
2.3 Model for the multi-period simulation	3
3 Implementation	5
4 Simulator interface	6
4.1 Single-period simulation	7
4.2 Multi-period simulation	10
5 Feedback and Data collection	11
6 Compatibility	11
7 Discussion and conclusions	12
References	13
Appendix A	15
A Software documentation	15
A.1 Setting up	15
A.2 Changing parameters	16
Appendix B	16
B Single-round simulation's instructions	16
Appendix C	16
C Multi-round simulation's instructions	16

Symbols and abbreviations

Symbols

B	Budget for a single period.
c^F	Cost of funding a project for one period.
c^R	Cost of re-evaluating a project.
m	Number of projects introduced in each period.
V_i	The random value of project i such that $V_i \sim N(\mu, \sigma^2)$.
v_i	The value of project i .
μ	Mean of the prior distribution for the project's values.
σ	Standard deviation of the prior distribution for the projects' values.
V_i^e	The random initial value estimate of project i such that $V_i^e \sim N(v_i, \sigma_e^2)$.
v_i^e	Initial value estimate for project i .
σ_e	Standard deviation of the initial estimates.
V_i^r	The random additional value estimate of project i such that $V_i^r \sim N(v_i, \sigma_r^2)$.
v_{ib}^r	The realized additional value estimate for project i .
σ_r	Standard deviation of the additional value estimates.
z_b^R	A binary decision vector used to denote which projects are re-evaluated in batch b .
z^F	A binary decision vector used to denote which projects are fully funded.
v_i^t	The value of project i introduced in period t .
s	The fraction of a project's value that is returned if the project is only partially funded.
$z^{FF,t}$	A binary decision vector used to denote which projects are fully funded in period t .
$z^{A,t}$	A binary decision vector used to denote which projects are conditionally funded in period t but <i>not funded in the next period</i> .
$z^{C,t}$	A binary decision vector used to denote which projects are selected to be conditionally funded in period t such that funding is continued in the next period.
v_i^{EB}	The latest available estimate for project i .

Abbreviations

DM	Decision maker
EV	Final portfolio's estimated value
HTTPS	Hypertext Transfer Protocol Secure
URL	Uniform Resource Identifier

1 Introduction

Firms, organizations, governments, and individuals seek to create value by allocating resources to different kind of projects. For example, firms do R&D projects to develop products and practices which will generate profits and governments need to decide how to allocate their healthcare budget.

It is generally difficult to know beforehand which projects will yield the most value ex post. Thus, the projects are selected based on uncertain ex ante estimates about these values. Because of the estimation uncertainties, the selected portfolio might be suboptimal. In particular, if funding is granted for a long period of time, resources may be tied to projects which later turn out to be unsuccessful [1].

The uncertainty in the projects' value estimates can be reduced by e.g., (i) purchasing additional evaluations for some projects, or (ii) by granting funding for some projects initially for only a part of their duration such that the continuation of this funding is based on an interim evaluation [1]. Using such means for reducing estimation uncertainties requires making trade-offs between (i) obtaining more accurate information about the projects, and (ii) using resources for acquiring additional value estimates and/or for funding projects that are discontinued so that only a fraction of the value of these projects can be retrieved. Methods of portfolio decision analysis have been developed to help the decision-maker (DM) to make these trade-offs in a cost-efficient way [1-3].

In this study we present a simulator software to explore how the DMs allocate resources between project funding and evaluation costs without recourse to formal methods. Moreover, the software can be used for conducting behavioural experiments on project selection, and as a management simulator tool for helping the DMs observe the consequences of their selection decisions. The software includes (i) a single-period simulator in which the DM needs to select a fixed number of projects but can purchase additional estimates about the projects' values prior to the selection decision, and (ii) a multi-period simulator, in which the DM is faced by a new set of project proposals in each period such that these projects can either be funded for their entire duration of two periods, funded for one period and then continued or abandoned based on interim estimates acquired at a cost, or rejected from the outset.

The rest of this study is structured as follows. Section 2 presents the model and Section 3 discusses the implementation of the tool. In Section 4 we describe how the subjects interact with the simulator. Section 5 explains what kind of data is collected from the simulations. Section 6 describes which browsers are compatible with the software and, finally, Section 7 concludes.

2 Project funding model

The following project funding model in the software is based on the model proposed in [1]. The DM is presented with a set of projects and she has three options to interact with them: The DM can fund projects, acquire better estimates about the

project's values or abandon projects. In this chapter we introduce the mathematical framework behind both simulations, single-period simulation and multi-period simulation.

2.1 Value model

Consider a set of project proposals $i = 1, \dots, m$, which if fully funded will yield values $v = [v_1, \dots, v_m]'$. These realizations are independent, identically distributed and normally distributed random variables $V_i \sim N(\mu, \sigma^2)$, where μ and σ^2 are known. Normal distribution was chosen because it is relatively easy for decision makers to understand. In a real applications, any distribution could be used.

At the time of the project funding decision, the DM does not know the values v but observes the estimates $v^e = [v_1^e, \dots, v_m^e]'$. Estimate for project i is a realization of the random variable V_i^e , which is a sum of the actual v_i and an error term ϵ_i : $V_i^e = v_i + \epsilon_i$, where $\epsilon_i \sim N(0, \sigma_e^2)$. The parameter σ_e is the known standard deviation of the initial estimate. More accurate estimates can be acquired at a cost of c^R . These additional estimates are realizations of $V_i^r = v_i + \epsilon_i^r$, where $\epsilon_i^r \sim N(0, \sigma_r^2)$, such that σ_r is known. Figure 1 illustrates the distributions of the the initial and additional value estimate acquired for a project whose actual value is $v_i = 10$. This has been generated using distribution $N(0, 10^2)$. For initial estimates we have used $\sigma_e = 10$ (Figure 1a) and for additional evaluations $\sigma_r = 10$ (Figure 1b).

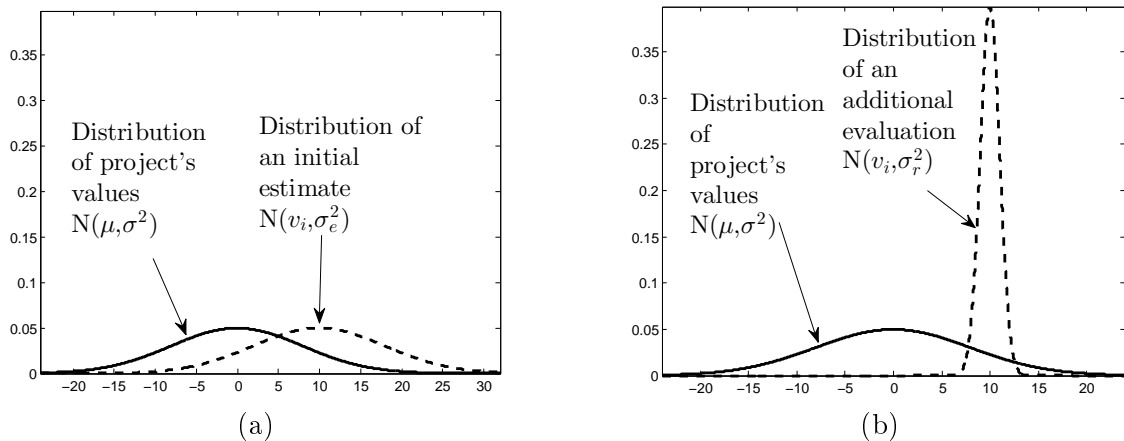


Figure 1: Distribution of the projects' actual values, and the initial and additional value estimates when the actual value is $v_i = 10$.

2.2 Model for the single-period simulation

In the single-period simulation model, the DM is presented with $i = 1, \dots, m$ project proposals, each of which lasts for one period. Before making her final funding decision, she is able to acquire additional value estimates v_{ib}^r for project batches, where b is the index for the batch. These additional estimates can be acquired for three batches in total and, thus, the DM can get at most four estimates (one initial

and three additional estimates) per project. The batch in each evaluation round can consist of any combination of the m projects. However, one additional evaluation for one project costs c^R . This cost is subtracted from the value of the final portfolio. The evaluation decision for each project i in batch b can be denoted by a binary vector $z_b^R = [z_{b,1}^R, \dots, z_{b,m}^R]$, where $z_{b,i}^R = 1$ if and only if the project i is evaluated in batch b . For each project i selected for re-evaluation, the DM observes a realization of random variable V_i^R . After she has conducted all of three additional evaluation batches or she does not want to evaluate any more projects, the DM can select which projects to fund. A funding decision is denoted by a binary decision vector $z^F = [z_1^F, \dots, z_m^F]$, where $z_i^F = 1$ if the project i is funded and $z_i^F = 0$ otherwise. Funding a project will yield a value v_i , while rejecting ($z_i^F = 0$) a project will not yield any value. Project funding is constrained by a budget

$$\sum_{i=1}^m z_i^F c^F \leq B, \quad (1)$$

where c^F is a project's fixed funding cost and B is the budget. The DM is trying to optimize her final portfolio value:

$$\sum_{i=1}^m \left(z_i^F v_i - \sum_{b=1}^4 c^R z_{b,i}^R \right). \quad (2)$$

The complete decision process in the single-period simulation is as follows:

1. The DM observes all the initial estimates v^e and is able to re-evaluate any number of projects as a batch process.
2. For each re-evaluated project i , the DM observes an additional estimate v_{ib}^r . If the DM is satisfied with the evaluations or has used up her three evaluation batches, the DM continues to the next step, otherwise the DM returns to the previous step along with the information gain of her additional evaluations.
3. The DM selects which of the projects to fund based on the evaluations.

Optimal strategies for similar project selection process settings are discussed in detail in [1].

2.3 Model for the multi-period simulation

The most noticeable difference between the simulation models is that in the single-period simulation the DM is given all the projects m at once, while in the multi-period simulation the DM is faced by m new project proposals in periods $t \in [1, T - 1]$, where T is the number of periods in the simulation. Like in single-period simulation, cost of re-evaluating a project is c^R . However, this cost is used in the budget constraint and is not subtracted from the value of the portfolio. Each project in the multi-period simulation last for two periods. Let us denote the values

of the projects of period t by v_i^t . A project has to be funded for two periods to yield a value v_i^t . If a project is funded only for a single period, the project will yield a scrap value sv_i^t . The parameter $s \in [0, 1]$ denotes how big a fraction of the project's value is retrieved if the project is funded only for a single period. A funding decision can be denoted by a binary decision vector $z^{FF,t} = [z_1^{FF,t}, \dots, z_m^{FF,t}]$, where $z_i^{FF,t} = 1$ if project i has been granted a full funding for two periods in the period t , $z_i^{FF,t} = 0$ otherwise.

Projects can be conditionally funded for one period and then re-evaluated or abandoned in the beginning of the next period. This incurs a cost of c^F in period t and c^R in period $t + 1$ per project. By doing this, the DM is able to observe additional realizations v_i^r of the random variable V_i^r for every evaluated project i . Only one additional estimate v_i^r can be obtained for each project as the projects last only for two periods. Conditional funding decisions are denoted by two binary vectors $z^{C,t} = [z_1^{C,t}, \dots, z_m^{C,t}]$ and $z^{A,t} = [z_1^{A,t}, \dots, z_m^{A,t}]$:

1. If the DM has conditionally funded a project but *has not continued* the funding in the next period, then $z_i^{A,t} = 1$.
2. If the DM has conditionally funded a project and *has continued* the funding in the next period, then $z_i^{C,t} = 1$.

Thus, if both $z_i^{C,t} = 0$ and $z_i^{A,t} = 0$ then the DM has not conditionally funded project i in period t .

In each period the DM has a budget of B for evaluation and funding. The budget constraint limits the DM's decisions in period t as follows:

$$\sum_{i=1}^m [c^F(z_i^{FF,t} + z_i^{A,t} + z_i^{C,t} + z_i^{FF,t-1} + z_i^{C,t-1}) + c^R(z_i^{C,t-1} + z_i^{A,t-1})] \leq B. \quad (3)$$

The DM is also restricted of evaluating and funding project i in the same period t . That is only one of the decision variables ($z_i^{FF,t}$, $z_i^{A,t}$, and $z_i^{C,t}$) can be 1:

$$z_i^{FF,t} + z_i^{A,t} + z_i^{C,t} \leq 1 \quad \forall i, t \quad (4)$$

The value of the portfolio in a T -period simulation is

$$\sum_{t=1}^{T-1} \sum_{i=1}^m v_i^t (z_i^{FF,t} + z_i^{C,t} + sz_i^{A,t}). \quad (5)$$

The portfolio selection process for the multi-period simulation is as follows:

1. The DM observes initial estimates v^e of m projects and is able to either fund (for two periods), re-evaluate (fund for one period and evaluate in the beginning of the next period) or reject each project.
2. For each project i which was evaluated in previous period ($t - 1$), the DM observes a new estimate v_i^r . The DM selects which of these re-evaluated projects

to fund (for one period) or reject based on both estimates. Simulator also presents m new projects and their initial estimates v^e . The DM is able to either fund (for two periods), re-evaluate (fund for one period and evaluate in the beginning of the next period) or reject each new project as previously. If the DM funded projects for two periods in the previous period ($t - 1$), the cost of funding these projects for one period (c^F per project) are subtracted from the DM's budget in this period as well. Also evaluation costs (c^R per project) of the previous period ($t - 1$) are subtracted from this period's budget.

3. The projects introduced two periods ago will expire, and are removed from the simulation. The DM will receive profits of these projects in the end of the simulation. The simulation repeats from step 2 until T periods have passed.

3 Implementation

The software is implemented in Google's App Script [9]. App script was chosen as a platform because it makes it possible to build simple web applications easily. Yet the engine is very limited in terms of visualizations and efficiency. App Script only supports a very small set of different kind of visualizations and they cannot be readily customized [9]. The implementation turned out to be quite slow (it takes about 1-3 seconds to load each page). The implementation is also highly dependent on the performance of Google's cloud service. The possible aggravation due to slow loading times was alleviated by adding a small loading icon each time the simulation is loading. This indicates to the DM that something is happening.

The software uses the ziggurat algorithm [4] to sample normally distributed pseudo-random numbers. The ziggurat algorithm belongs to the class of rejection sampling algorithms. Often pseudo-random number generators in video games are seeded with the current computer time. However, to facilitate comparisons between the input provided by different DMs, this simulator uses a fixed seed number. This means that the DM's will observe the same random numbers. This common random number method also reduces the variance of the results [15].

The default parameters for the simulator were chosen by experimenting. Parameters were chosen so that the optimal strategy would not be an extreme one: always funding the top projects, or always evaluating all projects and then funding the best ones. This was achieved by having a high standard deviation for the initial estimates and a small low standard deviation for the additional estimates, while having a budget constraint limiting the count of projects to six. This implies that the value gained by selecting better projects based on the additional, more accurate value estimates is likely to offset the costs of acquiring such estimates. Projects' values were also chosen in a such way that it does not always make sense to use the whole budget to fund projects, as they can yield negative value. The simulator was built to support a few different kinds of simulations just by changing the simulator's parameters. Table 1 shows the default parameters for both simulations and also the possible parameter options.

Table 1: Parameters of the two simulations and other options

Name	Symbol	Single-period	Multi-period
Mean of project's values	μ	0	0
Std of the projects' values	σ	10000	10000
Number of projects	m	10	8
Std of the init. est.	σ_e	10000	10000
Std of the add. est.	σ_r	8000	2000
Funding cost/project/period	c^F	5000	5000
Evaluation cost/project	c^e	700	1000
Budget	B	30000	30000
Number of periods	T	1	10
Number of repeats	T_r	10	1
Scrap value fraction	s	0	0
Skip simulation	<i>skip</i>	false	false
Number of training periods	T_t	3	NA

4 Simulator interface

The simulator consists of different views, which are shown to user in sequence. These views are the following:

1. General experiment instructions and log in
2. Instructions for the single-period simulation
3. Project selection view for the single-period simulation
 - (a) Evaluation and funding
 - (b) Results
4. Feedback for the single-period simulation
5. Instructions for the multi-period simulation
6. Project selection view for the multi-period simulation
7. Feedback for the multi-period simulation
8. Thank you screen

Simulation begins with the instructions, followed by the actual simulation and then by asking feedback about the simulation. It is possible to set the parameters of the simulator to skip either of the simulations. Skipping a simulation skips the simulation specific instructions, as well as the project selection and the feedback views.

4.1 Single-period simulation

The subject starts the simulation by reading the instructions and by entering her name in the input field. Other fields, such as sex and title, could easily be added to this view as well if needed. Instructions at this point explain what the simulation is about, but do not go into specifics. After the login screen, the subject is presented with the instructions for the single-period simulation. To avoid anchoring, no example strategies are presented in the instructions. This is to ensure that the subject does not anchor herself to the strategy presented in example, but figures out her own strategy. However, the instructions explain the parameters, the model and the objective for the subject. Instructions presented for the subjects are available as attachments B and C for both simulations.

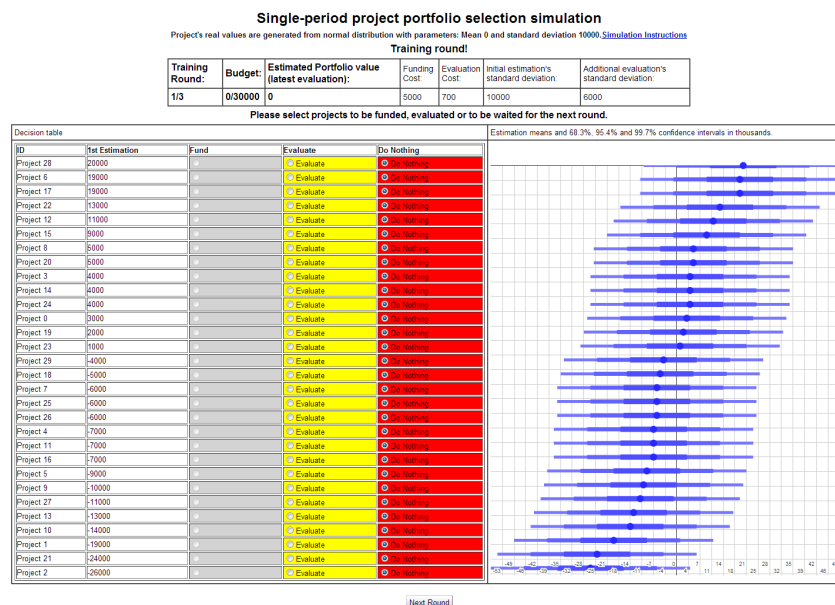


Figure 2: The initial view for the single-period simulation.

The first three rounds (by default) of the single-period simulation are training rounds. The subject is meant to learn the core mechanics of the simulation. These core mechanics include things like how to select projects using the interface and how uncertain the initial estimates are. These training rounds function exactly as real rounds of the single-period simulation, but they do not contribute to the final score presented in the end. Thus, the subject is able to experiment different strategies during the first few rounds.

The interface shows all the required parameters, and also an URL to the simulation instructions so that the subject is able to reread the instructions while playing. The subject is presented with m projects, their estimated values v^e , and a visualization showing these estimates and their confidence intervals. Projects are sorted in descending order based on the initial the estimates v^e . Figure 2 is a screen shot of the simulation at this stage.

The interface displays the estimated value (EV) of the subject's final portfolio.

This is calculated from the subject's best available estimate v_i^{EB} (the subject's latest evaluation for the project i) as follows.

$$EV = \sum_{i=1}^m \left(z_i^F v_i^{EB} - \sum_{b=1}^4 c^R z_{b,i}^R \right) \quad (6)$$

Confidence intervals of the estimates are shown in the visualization to guide the subject in interpreting the reliability of the estimates [10]. This visualization displays confidence intervals 68.3%, 95.4% and 99.7% (corresponding to $\pm\sigma_e$, $\pm2\sigma_e$ and $\pm3\sigma_e$, respectively) as a discrete color gradient. The subject is also able to see the exact values of the intervals' end points by placing the cursor over them.

As can be seen in Figure 2, the subject has currently two options per project to choose from. An additional third option, funding, becomes available after the batch evaluations have been conducted. These radio buttons are coloured in a way to indicate funding as a positive choice (green), evaluating as an uncertain choice (yellow) and rejection and no action as a negative choice (red). This color coding also enables the subject to make faster decisions and to make fewer errors with project selection [14]. To keep colour-blinds in mind, buttons are labelled, and they also appear consistently in the same order for all projects.

The subject is able to buy additional evaluations as batches at most three times. After each evaluation batch, the subject is able to see additional estimates for the projects. These new estimates appear similarly on the confidence interval visualization, as can be seen in Figure 3. The new estimates and their confidence intervals are visualized in red on top of the initial estimates, which are visualized in blue. This blue, red and white visualization should still be readable for people with Protanomaly (red weakness) [6], Deuteranopia (green blindness) [7] and Tritanopia (blue-yellow blindness) [8] color-blindnesses [5].

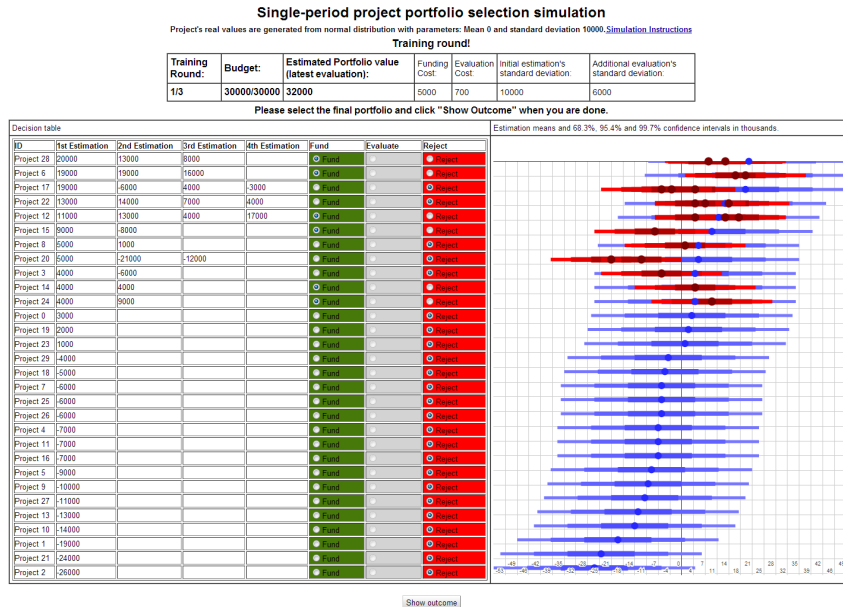


Figure 3: The funding phase for the single-period simulation.

After the subject has finished the re-evaluation process, the subject can continue to the next phase of project funding with a press of a button. The subject then selects projects to fund. After the subject has made her final choice, she clicks "Show outcome" button to proceed. After this, the simulator displays the actual values of the projects (v_i) in the selection table as well as in the confidence interval visualization as shown in Figure 4. The interface also displays the final value of the selected portfolio. The subject is able to compare this value to its estimated value. Because those projects whose values have been overestimated are more likely to have been selected than those with underestimated values, the estimated portfolio value is likely to be higher than its actual value. Hence, the subject is expected to experience post decision disappointment [1]. The subject is also able to compare the same information in project level by comparing the estimated and realized values of the projects in the selection table.

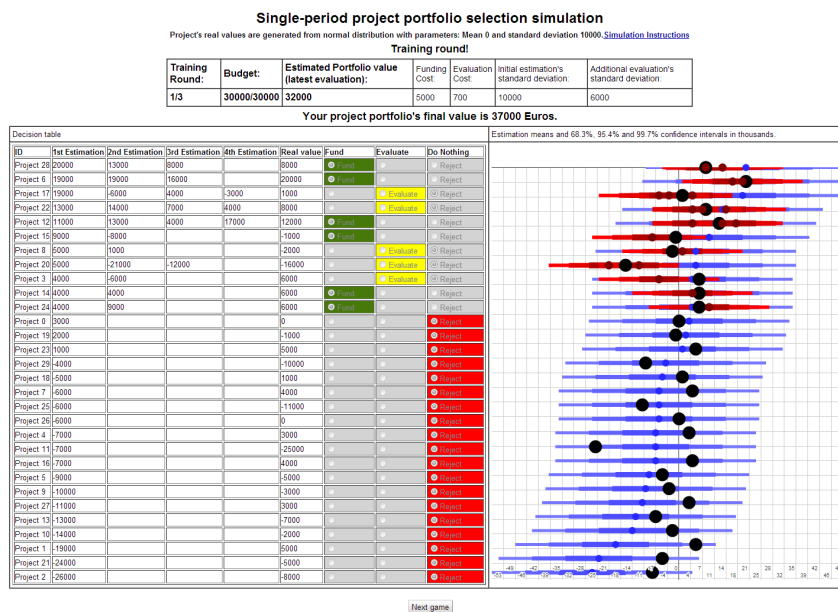


Figure 4: The result screen for the single-period simulation.

The subject is not shown any benchmarking information on her performance. For instance, displaying the value of the portfolio consisting of the best projects based on the initial value estimates alone (i.e, without acquiring additional estimates) might direct the subject towards adopting such a strategy; especially if this strategy yields higher value than her own strategy in the initial rounds. Also, using the optimal portfolio value as benchmark may have demoralizing effects, because the value of any portfolio selected by the subject is at most equal to this optimal value, and strictly lower for any strategy in which additional evaluations are acquired. It might be beneficial to show an average score of previous subjects and compare the current subject's score to it. However, this information is not available at this point of time.

4.2 Multi-period simulation

After the subject has finished all the repetitions of the single-period simulation, she will enter the multi-period simulation. The multi-period simulation works very similarly to the single-period simulation. However, the DM now has a budget for each period, there are more periods, and new projects are introduced in each period as discussed in Section 2.3. The initial setting of the multi-period simulation can be seen in Figure 5.

Evaluation and funding costs are also split differently compared to the single-period simulation: funding a project for two periods consumes resources both in the period in which the project is launched and in the following period. Evaluation cost is also now taken from the next period's budget and not from the current period's budget.

New column "Periods to go" is introduced to distinguish new projects from the projects launched in previous period. Projects are first sorted by "Periods to go" column, and then by their initial value estimates as illustrated in Figure 6. The intention with this is to help the subject concentrate on these two project groups separately (i.e. first funding the promising projects from the previous round and then looking which projects to re-evaluate from the set of recently launched projects). After a set of projects has been completed, these projects are removed from the simulation and they are considered either funded or rejected.

Estimated value is calculated in a different way than in the single-period simulation. This is because the multi-period simulation calculates the subject's final portfolio value differently: additional evaluations do not affect the final value and there are more rounds. The estimated value of the subject's portfolio EV can be formulated as follows:

$$EV_t = \sum_{i=1}^m \left(z_i^{FF,t} v_i^{EB,t} \right) + EV_{t-1} \quad (7)$$

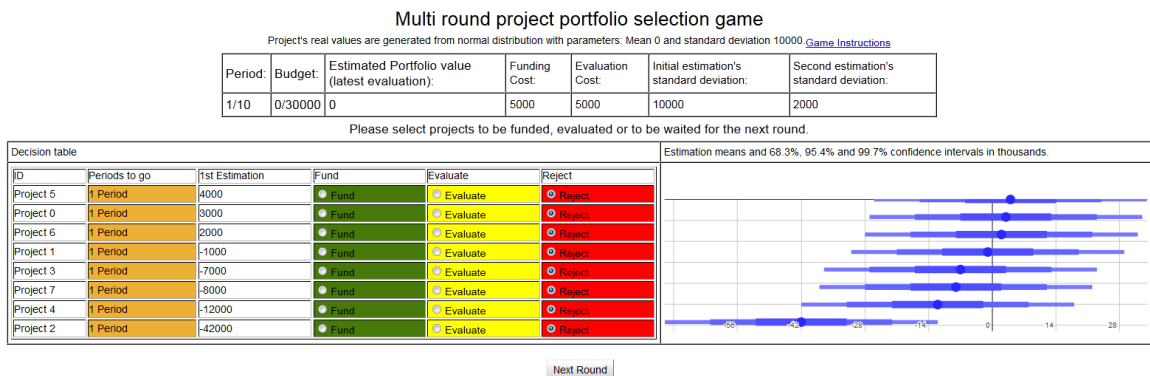


Figure 5: Multi-period simulation's first round.

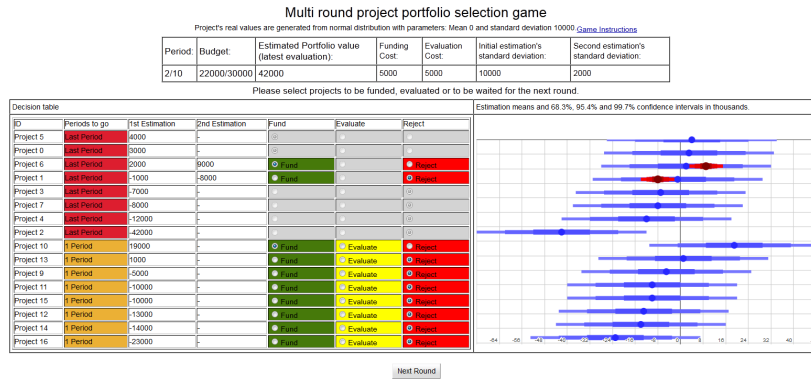


Figure 6: Multi-period simulation’s second round.

5 Feedback and Data collection

Each of the simulations requests feedback from the subject after she has completed either of the simulations. The feedback form asks the subject to write about the learning experiences she had as well as the kind of strategies the subject used during the experiment.

The simulation input and the feedback provided by the subjects are recorded to a Google Spreadsheet. The simulator stores information about which of the simulations were used (the single-period simulation or the multi-period simulation), time stamp (in human readable format and in milliseconds), the subjects identification code (subject’s name + random identification code), subject’s selections, the projects’ estimated and actual values, money spent, budget, the subject’s score in each period and in the whole simulation. The reason for using a random identification code is to prevent name clashes. The identification code is given when the subject has entered her name in the login screen, and stays unchanged during the simulation.

The subjects’ selections and the projects’ actual values are stored as matrices. In the selection matrix each period is separated by parentheses. The selection matrix uses number coding to distinguish between subjects’ different actions. Ones represent a funding choice, twos represent rejections and higher numbers present the additional evaluations. In the value matrix, projects’ actual values are on the first row and their estimated values on the following rows in the same order as they appear in the simulation.

6 Compatibility

Compatibility of the software was tested using three different browsers: Internet Explorer 10, Chrome 30.0.1599.101 and Firefox 25.0. Simulator runs on each browser with somewhat varied performance and usability. Each browser displays the interface elements a bit differently. For example, Chrome displays texts of disabled text boxes (texts of the instruction screens) in a brighter color than Firefox does. This is not really an issue as elements are still placed correctly, and their visual appearance

is acceptable. However, some issues arise when the browser's native zoom is used. While the zoom works perfectly in Chrome, Firefox is not entirely able to display the visualization correctly as can be seen in Figure 7.

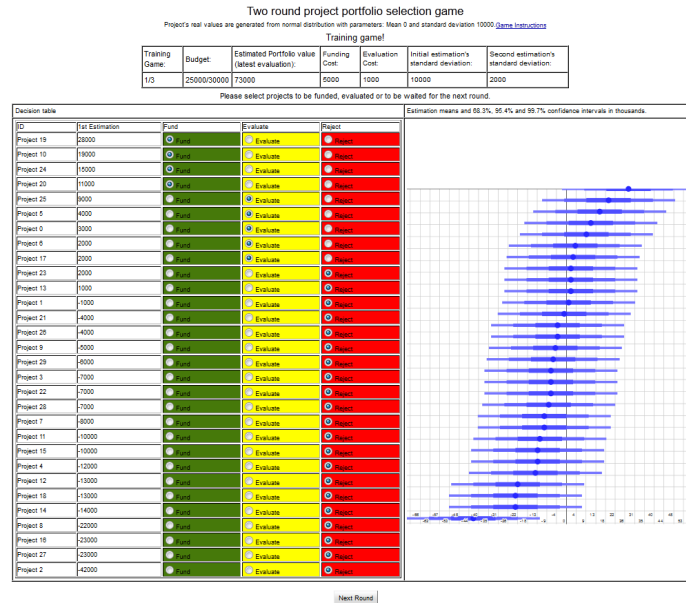


Figure 7: Visualization is not displayed correctly on Firefox when Firefox's zoom is used.

The software runs decently in all of the tested browsers except in Internet Explorer. Internet Explorer displays a security warning pop-up when the site is opened. This is because the site is not using HTTPS protocol. If the subject clicks "yes" in this pop-up, the site stops functioning. Internet Explorer also complains that scripts on the site might be slow. Again, if the subject clicks "yes" the site stops functioning. On the other hand, clicking "no" just closes the pop-up for a few seconds until it pops up again. It is thus advisable that Internet Explorer users disable this pop-up before starting the simulation, as it slows down the simulator significantly.

While the simulator functions in all tested browsers, it functions best on Chrome. This is not surprising as Chrome is Google's own technology like App Script. The performance is the worst with Internet Explorer, because of the above error messages. Thus, if possible, the simulator should be used with Chrome, but if this is not possible, Firefox functions as well.

7 Discussion and conclusions

Our software provides a platform to conduct project selection experiments. Software implements two types of project selection simulations. These two types of simulations are based on theoretical results developed in [1] and [3]. The first experiments with the software are meant to be exploratory. The software can be used for illustrating the phenomenon at post-decision disappointment, and informing subjects about the consequences of selecting projects based on very uncertain estimates.

By recording the subjects' input, the software can be used to analyse various aspects of the subjects' decision making processes; for instance, how long it takes for the subjects to conduct each part of the experiment, what kinds of strategies are the subjects using, or how their speed and score improve over time. Results could also be used to calculate simulators learning curve: How the subject's performance can be expected to improve over time [12]. It would also be informative to know how strong a role anchoring (a judgement bias in which decision makers are systematically influenced by uninformative and random starting points [11]) plays in portfolio selection. It is also interesting to see whether the subject's abandon projects in the current period in hopes of getting to select projects with higher values in the next period in the multi-period simulation [13].

Experimenting with different kinds of simulation settings could provide useful information: transferring unspent money from previous period to current period, providing benchmarking information, asking the subjects to seek the highest profit projects instead of maximizing the value of the whole portfolio, varying funding costs and partially funding projects. Experimenting with subjects with different kinds of background (such as different degrees) and seeing how the decisions of such groups differ, could also provide interesting information. Nevertheless, carrying out experiments using the software in the form that is reported in this study would (i) provide a base case for developing hypotheses to be tested in subsequent experiments, and (ii) indicate which changes in the simulation parameters, in the model, or in the implementation would be the most relevant. The actual experiment and proposed simulation variations remains as future work.

References

- [1] Vilkkumaa, E., Liesiö, J., Salo, A. 2014. *Optimal Strategies for Selecting Project Portfolios Using Uncertain Value Estimates*. European Journal of Operational Research, 223, 772–783.
- [2] Salo, A., Keisler, J., Morton, A. 2011. *An Invitation to Portfolio Decision Analysis*. In: *Portfolio Decision Analysis: Improved Methods for Resource Allocation*. Salo A., Keisler, J., Morton, A. Springer, New York.
- [3] Vilkkumaa, E., Salo, A., Liesiö J., Siddiqui, A. 2013. *Optimal Policies for Funding Breakthrough Projects*. Manuscript.
- [4] Marsaglia, G., Tsang, W. 2000. *The Ziggurat Method for Generating Random Variables*. The Florida State University and University of Hong Kong.
- [5] Hess, R. 2000. *Can Color-Blind Users See your Site*. Microsoft Developer Network. Updated 09.10.2000. Referenced 30.10.2013. Available: [http://msdn.microsoft.com/en-us/library/bb263953\(v=vs.85\).aspx](http://msdn.microsoft.com/en-us/library/bb263953(v=vs.85).aspx).
- [6] Yasuma, T., Tokuda, H., Ichikawa, H. 1984. *Abnormalities of Cone Photopigments in Genetic Carriers of Protanomaly*. Arch Ophthalmol.

- [7] Ishihara, S. 1972. *Tests for Colour-Blindness*. Kanhara shuppan Co. Ltd.
- [8] Wright, W. 1952. *The Characteristics of Tritanopia*. The Journal of the Optical Society of America, *42*, 509-517.
- [9] *Google Apps Script Documentation*. 2013. Google Ltd. Referenced 16.11.2013. Updated 05.11.2013. Available: <https://developers.google.com/apps-script/>.
- [10] Masson, M.E., Loftus, J., Geoffrey, R. 2003. *Using Confidence Intervals for Graphically Based Data Interpretation*. Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, *57*, 203-220.
- [11] Gretchen, B., Chapman, E., Johnson, J. 1999. *Anchoring, Activation, and the Construction of Values*. Organizational Behaviour and Human Decision Processes, *79*, 115–153.
- [12] Yelle, L.E. 1979. *The Learning Curve: Historical Review and Comprehensive Survey*. Decision Sciences, *10*, 302–328.
- [13] Dixit, A.K., Pindyck, R.S. 1994. *Investment Under Uncertainty*. Princeton Press, Princeton.
- [14] Green, B.F., Anderson, L.K. 1956. *Color Coding in a Visual Search Task*. Journal of Experimental Psychology, *51*, 19-24.
- [15] Averill, M., Law, W., Kelton, D. 2000. *Simulation Modeling and Analysis, Third edition*. McGraw-Hill Higher Education, New York.

A Software documentation

The simulator consists of four Apps Script files: Code.gs, randomGen.gs, timenow.gs and parameters.gs. Code.gs contains most of the simulator code, randomGen.gs contains a random number generator for normal distribution, timenow.gs has a function to get the current time, and parameters.gs contains the simulator's parameters. All these files have to be located in the same project in the script manager.

The simulator starts from doGet() function. This function is called when a user opens the site. GUI works as creating and deleting Google App Script elements. When a new page is created, previous elements are removed and new elements are added. All the simulator's files and functions are documented.

A.1 Setting up

To set-up the simulator, one has to create a new Google Spreadsheet and copy the code files to the script manager. Script manager can be found under Tools. Create a new project and one file for each code file and move contents to scripts sheets on Google Scripts editor.

After all files has been copied successfully, software has to be authorized. Click run, select "handlerFunction" and give permissions to the application (do not care about the error messages on running the handlerFunction). After this one needs to change the location where the simulator's period logs are stored. Logs are stored in a Google Spreadsheet identified by "SheetId" parameter, which is located in the parameters file. Open this file and change the variable "SheetId" on top of the file to the sheet identification code of the new sheet. Identification code of a Google Spreadsheet can be found from the spreadsheets url after "?key=" and before "&" sign. For example original projects sheet ID (0Akt6cTpecDdHdJR256dDgzUkZnaHJaMWZtQ nVPYUE) can be found from URL of a sheet: https://docs.google.com/a/avaus.fi/spreadsheet/ccc?key=0Akt6cTpecDdHdJR256dDgzUkZnaHJaMWZtQnVPYUE&usp=drive_web#gid=0. Authorize the script files with your own Google initials, so that the use of the simulator can use your credentials when writing to the data file.

After all this, the simulation should be functional. You may test the simulation by clicking "Publish" on the spreadsheet, "Deploy as web app" and copying the url given in the "Current web app URL". Paste this URL into a browser and you should be able to test the simulator. This is the simulator's test environment and this site is updated every time the script files are saved. This url should not be given away for end users. To publish the current version of the simulation to the end users, one needs create new version of the web application. This is achieved by clicking "File" and then "Manage versions". Type something in the text field and click "Save new version". Now a new publishable version of the software is created. This version can be published from "Deploy as web app" window (discussed earlier). Select the newest project version, execute the application with your credentials, and give everyone access to this application. Click update when you are done. Now a window should appear which will show you an URL to the web application available

for everyone.

A.2 Changing parameters

Simulator's parameters can be changed from parameters.gs file. All the parameters are commented and should be quite self-explanatory. The simulator does not support project simulations with multiple different kinds of parameters to run at once (e.g. making the simulators to run in separate URL addresses). However, this can be solved by creating a new project (new spreadsheet) and performing the set up described in the above section again. The two current simulators can be separated by setting parameters of one project to "var SkipGames=[0,1]" and other to "var SkipGames=[1,0]".

B Single-round simulation's instructions

Your task is to select investment projects based on estimates about their future value. You have a budget of 30 000 euros that can be used to fund projects. You are also able to obtain additional, more accurate estimates about the projects' values. This can be done at most three times for each project. The cost of funding a project is 5000 euros. Your goal is to maximize total value of your funded projects. In addition, 700 euros is subtracted for each additional estimate acquired. The selection decision is done in two stages:

1) Based on the initial value estimates, select which projects are re-evaluated as a batch process. You are able to conduct at most three rounds of these batch evaluations.

2) After obtaining the additional value estimates (three times or less), select which of the projects to fund.

There are 30 project proposals in total. The values v of the projects are normally distributed with a zero mean and a standard deviation of 10000 euros. The estimates are obtained from these values by adding a normally distributed error term eps1 or eps2 with a zero mean and a standard deviation of 10000 euros (initial estimate $vE_initial$) or 6000 euros (additional estimate $vE_additional$).

$$v \sim N(0,10000)$$

$$vE_initial = v + \text{eps1}, \text{eps1} \sim N(0,10000)$$

$$vE_additional = v + \text{eps2}, \text{eps2} \sim N(0,6000)$$

This simulation is repeated ten times. Your final score will be calculated as the sum of the value of the portfolios of each simulation. The simulation starts with three training rounds, which are not included in the final score. Good luck!

C Multi-round simulation's instructions

Consider a multi-period setting, in which your task is to select investment projects based on estimates about their future value. In each period, you have 30 000 euros

that is to be divided between (i) project funding and (ii) costs of obtaining additional, more accurate estimates about the projects' values. The cost of funding a project for one period is 5000 euros, and the cost of obtaining an additional estimate for one project is 1000 euros. The duration of each project is two periods. The value of a project is realized only if it is funded for its entire duration. Your goal is to maximize total value of your funded projects (left over budget is not included in this value).

In each period, eight new project proposals become available. Based on the initial value estimates of these proposals, your task is to select

- which projects to fund for the entire 2 periods (cost 5000 euros now and 5000 euros in the next period), and
- which projects to fund for 1 period and to re-evaluate in the next period (cost 5000 euros now and 1000 euros in the next period).

Moreover, based on the additional estimates obtained for those projects that you decided in the previous period to fund for just 1 period, your task is to select

- which projects to continue funding for one more period (cost 5000 euros), and
- which to reject.

The values v of the projects are normally distributed with a zero mean and a standard deviation of 10 000 euros. The estimates are obtained from these values by adding a normally distributed error term eps1 or eps2 with a zero mean and a standard deviation of 8000 euros (initial estimate $vE_initial$) or 2000 euros (additional estimate $vE_additional$) - $v \sim N(0,10000)$

- $vE_initial=v+\text{eps1}$, $\text{eps1} \sim N(0,8000)$
- $vE_additional=v+\text{eps2}$, $\text{eps2} \sim N(0,2000)$

This simulation is run only once. Good luck!