Risk-informed project selection for developing breakthrough technologies

Konsta Parkkali

School of Science

Thesis submitted for examination for the degree of Master of Science in Technology.

Espoo 31.7.2024

Supervisor

Prof. Ahti Salo

Advisor

Prof. Ahti Salo

Aalto University
School of Science

Aalto University
School of Science

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi
Abstract of the master's thesis

Author Konsta Parkkali

Title Risk-informed project selection for developing breakthrough technologies

Degree programme Mathematics and Operations Research

Major Systems and Operations Research **Code of major** SCI3055

Supervisor Prof. Ahti Salo

Advisor Prof. Ahti Salo

Date 31.7.2024 **Number of pages** 51 **Language** English

Abstract

Public research funding programs advance innovative research and produce significant societal benefit through the development of breakthrough technologies. However, these programs have in general been criticized for inefficient allocation of funds, hampering the advancement of breakthroughs. Some of the hypothesized shortcomings of these programs include the absence of a holistic portfolio approach to risk management, excessively risk-averse selection of individual projects, and not funding projects conditionally in multiple stages.

In this thesis, we develop a stochastic optimization model for formulating risk-informed research funding policies to support the development of breakthrough technologies. We include an option to abandon projects in the model, granting the decision maker a possibility to experiment with a large set of projects by launching them for a set period of time, and committing resources only to those which hold the most promise after initial experimentation. Furthermore, we introduce concentration risk by allowing different project sizes and we model portfolio level risk using the Value-at-Risk framework.

Our numerical results are aligned with the aforementioned criticism towards research funding programs. The results show that there can be a clear trade-off between supporting the development of breakthrough technologies and research in a more general sense. In light of our numerical results, conditional project funding is important in risk-informed development of breakthrough technologies. However, conditional funding of small projects did not function as a risk mitigate due to the stiffness of the decision model. Rather, we found that more risk-seeking funding policies can be formed by increasing the duration and share of the conditional funding. Nevertheless, conditional funding of large projects can be a tool for risk management and promote the development of breakthrough technologies.

Keywords Project portfolio optimization, portfolio decision analysis, risk constrained optimization, Value-at-Risk, real option, stochastic optimization

Aalto University
School of Science

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi
Diplomityön tiivistelmä

Tekijä Konsta Parkkali

Työn nimi Riskitietoinen projektivalinta läpimurtoteknologioiden kehittämiseen

Koulutusohjelma Mathematics and Operations Research

Pääaine Systems and Operations Research **Pääaineen koodi** SCI3055

Työn valvoja Prof. Ahti Salo

Työn ohjaaja Prof. Ahti Salo

Päivämäärä 31.7.2024 **Sivumäärä** 51 **Kieli** Englanti

Tiivistelmä

Tutkimusrahoitusohjelmat edistävät innovatiivista tutkimusta ja tuottavat merkittävää yhteiskunnallista hyötyä läpimurtoteknologioiden kehityksen kautta. Näitä ohjelmia on kuitenkin yleisellä tasolla kritisoitu varojen tehottomasta allokoinnista. Tehottomuuden takana oleviin puutteisiin on arvioitu kuuluvan muun muassa portfolionäkökulman puuttuminen riskienhallinnasta, yksittäisten projektien ylivarovainen valinta, sekä projektien rahoittaminen ilman väliarvointeja.

Tässä työssä kehitetään stokastinen optimointimalli läpimurtoteknologioiden kehittämiseen tähtäävien riskitietoisten tutkimusrahoitustoimintaperiaatteiden muodostamiseksi. Malli sisältää reaalioption uudelleenarvioitujen projektien hylkäämiseksi, mikä antaa päätöksentekijälle mahdollisuuden käynnistää iso joukko projekteja kokeilujaksi ja siten sitoutua rahoittamaan vain lupaavimpia projekteja. Mallissa voidaan tarkastella eri kokoisista projekteista aiheutuvaa keskittymäriskiä, jota mallinnetaan portfoliotasolla Value-at-Risk kehikolla.

Työn numeeriset tulokset tukevat aiemmin esitettyä kritiikkiä tutkimusrahoitusohjelmia kohtaan. Tulokset osoittavat, että optimaalisten rahoitustoimintaperiaatteiden muodostamiseen vaikuttaa vahvasti se, tavoitteleeko päätöksentekijä läpimurtoteknologioiden kehittämistä vai tieteellisen tutkimuksen edistämistä yleisemmällä tasolla. Numeeristen tulosten valossa projektien ehdollinen rahoittaminen on tärkeä osa riskitietoista läpimurtoteknologioiden kehittämistä. Pienten projektien ehdollinen rahoittaminen ei kuitenkaan vähentänyt rahoitusohjelman riskitasoa mallin rajoitusten ja jäykkyyden takia. Päinvastoin, ehdollisen projektirahoituksen lisääminen ja pidentäminen nosti riskitasoa. Suurten projektien ehdollinen rahoittaminen osoittautui keinoksi hallita rahoitusohjelmien riskitasoa sekä tukea läpimurtoteknologioiden kehittämistä.

Avainsanat Projektiportfolio-optimointi, portfoliopäätösanalyysi, riskirajoitettu optimointi, Value-at-Risk, reaalioptio, stokastinen optimointi

Contents

1 Introduction

There is a widespread desire to foster innovative research, fueled by the anticipation of developing new breakthrough technologies. These technologies may be based on newly discovered physical or chemical phenomena, allowing the development of novel products with profound and irreversible impacts on society. An illustrative example of such groundbreaking discovery is the layered crystal structure utilized in the modern lithium-ion battery, which is a pivotal component in all modern portable consumer electronics and electric vehicles (Li et al., 2018).

Innovations leading to breakthrough technologies are exceedingly rare. Moreover, it is typically impossible to know whether an early-stage research and development (R&D) activity will result in such technologies. Hence, funding research in the hope for breakthrough technologies is an inherently risky activity. As most funded research fails to produce significant value, the justification for granting such funding can be challenging (Goldstein and Kearney, 2020). Consequently, private companies often avoid these risks by underinvesting in research and development of novel products and ideas (Griliches, 1991). In addition, companies can face immense pressure to generate profit in the short-term, while the development of new technologies is in general a time-consuming process. Thus, taking on an overly ambitious R&D activity and failing to demonstrate results in the short-term can result in losing the relevant stakeholders' trust and even lead to the company's downfall. Conversely, the willingness of a company to risk its short-term profit is found to be one of the key drivers for innovation (Tellis et al., 2009).

These issues emphasize the importance of public research funding. In fact, the tremendous positive impact of breakthrough technologies to national economic growth has driven the establishment of public research funding programs targeted at groundbreaking discoveries (Sharpe et al., 2013). However, more knowledge is needed to understand what kind of conditions and policies help make such discoveries (Grilli et al., 2018). Typically, funding programs attempt to recognise and invest in highly promising research activities, whether in public or privately owned research departments. For example, the Research Council of Finland has launched a research funding program BioFuture2025 to promote scientific breakthroughs for curbing climate change and the overuse of natural resources. This is pursued by seeking novel research ideas and opportunities, and by granting funding for ambitious and even risk-taking research activities (Research Council of Finland, 2017). Nevertheless, public research funding agencies still face the same challenges of balancing between potential high impact discoveries and risk of no return as do private companies (Goldstein and Kearney, 2020). They also suffer from the pressure of demonstrating short-term results discouraging the funding of risky research (Franzoni et al., 2022). There has been a growing criticism towards public research funding programs for their shortcomings in addressing these issues (Buzzacchi, 2022).

R&D activities are typically completed in the form of projects. Consequently, there exists many decision-making and management frameworks, not to mention full-blown optimization models for efficient R&D project selection and management (Elbok and Berrado, 2017; Gupta et al., 2022). Various organizations in the private sector have successfully implemented these sometimes very case specific and complicated models and frameworks. Furthermore, empirical evidence suggests that active management practices are linked to increased success of radical innovation (O'Connor et al., 2008). Yet, such direct project management approaches which are commonplace in the private sector are not usually seen in use in public research funding programs (Goldstein and Kearney, 2020). Rather, these programs tend to have minimal influence on the research projects after initial funding is granted.

There exists some empirical evidence of effective active project level risk management in public research funding programs. One example of such risk management is the real option of evaluating and abandoning projects, which can reduce the risk of committing resources to underperforming projects. This frees resources for other more promising projects. However, there is little information on the use of such real options as portfolio level risk mitigants in public research programs. (Goldstein and Kearney, 2020)

The aforementioned abandonment option has been studied extensively in frameworks where the goal is to maximize the expected value of a project portfolio and it has been showed to hold significant value in certain conditions (see e.g., Santiago and Vakili, 2005). By utilizing abandonment options, Vilkkumaa et al. (2015) explored the shaping of long-term funding policies with the aim of fostering breakthrough technologies. They found that when the objective was to maximize the number of breakthrough technologies in the long-term, the abandonment option did indeed have significant value. In these settings, an optimal policy was to launch a large number of projects but to abandon a high portion of them later. This policy has similarities to the "spray and pray" strategy familiar from the venture capital industry (Lerner and Nanda, 2020). However, the policy was accompanied by the trade-off of a decreased expected portfolio value when compared to a funding policy aimed purely at maximizing this expected value. This was partly caused by the cost of evaluating many projects and partly by the resources wasted on abandoned projects. Yet, even without resulting in breakthroughs, R&D projects can still produce significant value by merely contributing to the general scientific knowledge, which can be of use in future research. Thus, the outcomes stemming from such basic R&D activities also play an important role in fostering breakthrough technologies.

In settings such as those in Vilkkumaa et al. (2015) where the number of available projects is large and the projects are independent, risk is not of significant concern due to diversification benefits. However, this is not always the case. First, projects can be large, leading to fewer funded projects. Second, projects can have significant interdependencies. For example, very narrowly focused national funding policies have been gaining popularity in recent times (Janssen, 2019), which could result in funding highly correlated projects which all depend on the state of the underlying industry. However, funding agencies do not generally consider portfolio risk holistically when choosing research projects. Instead, the "one by one" method is typically used, i.e., projects are scored individually, and the best ones are chosen, which can lead to favouring less risky projects and aggravate the problem of being too risk-averse on the portfolio level (Franzoni et al., 2022).

In this thesis, we explore with a decision model how funding agencies could approach funding risky research projects to pursue breakthrough technologies whilst demonstrating short-term results. Furthermore, we study how active project management, such as the abandonment option, could be utilized to manage the portfolio level risk in the same context. We investigate these questions by extending the quantitative decision model of Vilkkumaa et al. (2015). We allow the decision maker to choose between funding projects with different risk levels, caused by the projects' varying sizes. We formulate a 3-stage stochastic optimization problem with the objective to maximize breakthrough technologies while using the Value-at-Risk framework to measure the short-term riskiness of the portfolio. We solve the Pareto front using the sample average approximation approach together with multi-objective optimization and analyze the Pareto efficient solutions both with and without an abandonment option. Furthermore, we analyze alternative project selection and management strategies, and in the process, we provide insight into key principles for shaping risk-informed funding policies for developing breakthrough technologies.

The rest of this thesis is organized as follows. In Chapter 2 we review the related literature covering topics such as portfolio decision analysis and risk constrained project portfolio optimization. In Chapter 3 we formulate the optimization model to guide the shaping of risk-informed funding policies. In Chapter 4 we present a method for solving the formulated optimization problem and derive needed analytic results. In Chapter 5 we analyze the numerical results and heuristic portfolio management strategies, the effectiveness of which are illustrated with Monte Carlo simulation. In Chapter 6 we discuss the results, consider ideas for future research, and conclude the findings.

2 Background

2.1 Portfolio decision analysis

The strategic allocation of resources to competing activities stands as one of the foremost considerations all organizations must address. Such activities are often completed in the form of projects, geared towards achieving the organizations' set objectives. A project can be defined as a transitory plan for reaching a preset goal and is typically accompanied by resource constraints such as budgets and deadlines. These constraints force organizations to pursue only a selected number of projects as there are usually more proposals for projects than available resources. Moreover, completing a project successfully and in a timely manner requires the ability to effectively manage these scarce resources in terms of coordinated efforts.

Portfolio management is the area of expertise concerned with the efficient allocation of resources across a collection of investments. In a holistic sense, portfolio management can be seen to cover all related aspects of this process such as the identification of different investments alternatives, the selection of the most promising ones, the distribution of resources to the ones selected, the monitoring of the investments, and all subsequent adjustments to the previous decisions (Salo et al., 2023). Like typical financial instruments such as stocks and bonds, projects can be viewed as investments from which the investor expects to obtain future benefits. As with practically all investments, these future benefits are uncertain as a project's outcome may fall short of initial expectations. Consequently, handling risks play a pivotal role in managing a portfolio of projects.

Portfolio management appears most often in the context of managing traditional financial portfolios consisting of stocks and bonds. However, the field of project portfolio management has emerged especially in contexts such as R&D, IT and construction (Elbok and Berrado, 2017). There are profound differences between these two contexts. To begin with, most projects are indivisible and must therefore be considered as lump investments. Moreover, projects are temporary endeavors which, in some cases, must be completed subject to time constraints.

Second, observing an ongoing project's status and estimating the project's future value is not trivial. It may require a tremendous amount of effort from a team of experts and multiple well-chosen criteria to understand the project and all the uncertainties affecting its future outlook (Rode et al., 2022). Moreover, as projects are unique, there is rarely sufficient historical data which could be used as a reference in decision making. In contrast, the prices of typical financial instruments are readily available in the stock market without cost, and they can be observed in an almost continuous manner. Furthermore, even though similar uncertainties hold as with projects, there exists a significant amount of historical data on financial asset prices which can be used together with well established statistical methods to estimate the behaviour and risk-return characteristics of future prices.

A mathematical approach to portfolio management emerged first in the 1950s with the introduction of the modern portfolio theory (Markowitz, 1952). The theory was based on the idea of maximizing the expected return of a portfolio of financial assets while having a quantifiable risk constraint. During the same period of time, the issue of effective project portfolio selection also emerged as an important decision problem with the need for systematic ways to evaluate and compare different project opportunities (Mottley and Newton, 1959). To tackle these issues, the discipline of portfolio decision analysis (PDA) started to form.

Defined formally, PDA is the discipline focusing on improving methods for selecting a subset of discrete objects from a wider set of alternative whilst taking into consideration the relevant uncertainties, constraints, and the decision maker's preferences (Salo et al., 2011). The tools employed in PDA often originate from the field of decision analysis, as PDA problems can be seen as extensions of conventional decision problems. In this context, the decision maker is allowed to choose not only on a single choice but on multiple selections from a large set of alternatives. This can often make the problem much more complicated, as aspects such as non-linear project dependencies and various uncertainties often require modelling on portfolio level (Micán et al., 2020). Furthermore, the curse of dimensionality emerges in multiselection decision problems. For example, there is only 100 ways to choose one project out of a set of 100 options. In contrast, when the decision maker is allowed to pick 5 projects, the number of distinct choices grows to $\binom{100}{5}$ or roughly 75 million. Thus, analysing all possible options one by one is often infeasible.

According to a recent review on PDA research, uncertainty and risk were identified as some of the key challenges in many PDA problems (Liesiö et al., 2021). The typical approaches for tackling uncertainties in PDA problems are based on classic probability theory and include methods such as decision trees and Monte Carlo simulation. However, approaches such as fuzzy logic and scenario modeling are also well established. In their review, Liesiö et al. (2021) also list out methods used for managing risk such as robust optimization, worst-case analysis, concave utility functions, and constraining risk-measures. Furthermore, they state that the availability of data or access to expert judgement are essential factors in determining to what extent uncertainties can be modelled. As the number of projects can be vast, the task of doing a detailed analysis, perhaps including a thorough probability assessment on each project can be overwhelming.

Another persistent challenge in many applications of PDA is that it can be hard to elicit the decision makers' preferences. There may be multiple decision makers who have different preferences for to the various objectives, including acceptable level of uncertainty. Typical methods for tackling these challenges are different preference elicitation techniques and multi-attribute value theory (see e.g., Gasparini et al., 2022). Often, the top management in organizations makes the final resource allocation decisions. Thus, to foster robust decision processes, it is important that the theories and methods used in PDA are transparent, understandable, and presentable. Furthermore, having a well-defined decision process makes structured evaluation of past decisions possible (Salo et al., 2011).

2.2 Related literature

2.2.1 Funding innovative research

Sharpe et al. (2013) characterize breakthrough technologies as revolutionary innovations based on new, underexplored, or untapped physical, chemical and biological phenomena, resulting in significant and permanent changes through enhancing existing products or giving rise to new ones. The concept of exceptionally excellent R&D results producing immense value is a more widespread phenomena and is encapsulated by varying terms such as radical, disruptive, or high-risk innovation (see e.g., Govindarajan and Kopalle, 2004). However, studies on these topics have mostly come in the form of empirical examinations.

Vilkkumaa et al. (2015) studied the fostering of breakthrough technologies by modelling research funding policies quantitatively. In their study, they examined the differences of optimal funding policies when the objective varied between maximizing the expected value of a project portfolio and the number of breakthrough technologies. They considered these technologies to stem from rare exceptionally excellent projects defined by an excellence threshold on the projects' ex-post values. They showed that there can be a significant trade-off between these two objectives, and that the real option of evaluating and abandoning projects is valuable especially in uncertain conditions where the values of the projects become more precise quickly after the initial launch. They propose that breakthrough technologies can be best fostered by experimenting with a large set of projects and committing resources only to the projects that seem promising after initial experimentation.

Klingebiel and Rammer (2011) studied the effects of breadth, uncertainty, and selectiveness in the context of promoting innovation using data from 1500 German companies. In-line with Vilkkumaa et al. (2015), they found that selecting a broad set of projects is particularly beneficial in uncertain conditions and with the option to abandon projects. The study of O'Connor et al. (2008) covered the management practices of large established firms by collecting data from 85 radical innovation efforts. They revealed that the use of real options such as the option to evaluate and abandon projects has a positive effect on radical innovation success, further supporting previous findings.

Goldstein and Kearney (2020) studied the project management approaches used in ARPA-E, a governmental research funding agency in the US. They found that real options, such as the options to abandon, contract, and expand project budgets and timelines were valuable especially when projects performed poorly, as this reduced the risk of committing funding to weakly performing projects. This suggests that active management could be used effectively in the context of project portfolio risk management. However, they did not analyze the portfolio effects of the real options.

Furthermore, they note that the management decisions on individual projects could be explained in more detail if considering the whole portfolio view.

The aforementioned findings are loosely in agreement with the general guidelines provided by Lerner (2009), who suggests that publicly funded programs supporting entrepreneurial activities should be sufficiently preserving while still holding the option to abandon failing activities. Too uncertain or short-term funding can, on the other hand, encourage risk-averse research and result in a loss of motivation (Manso, 2011; Heinze, 2008). This was also recognised by Franzoni et al. (2022) who examined reasons behind the bias against risky research. The authors hypothesize several reasons for this and propose ideas that support risk-taking research such as diverse funding boards, prizes for achieving breakthroughs, funding projects conditionally in multiple stages, and the inclusion of a portfolio view in research project portfolio selection.

Buzzacchi (2022) further discusses the inefficiencies in the allocation of funds by public research funding programs. The prevailing project selection strategy is the one by one approach, where project proposals are evaluated individually and the top-ranked ones are chosen until the budget is depleted. Funding agencies also tend to neglect the benefits of diversification, leading to insufficient support for risky research and an undesirably low level of risk aversion. Moreover, conditional project funding remains underutilized, even though it is widely adopted by private investors in the venture capital industry.

2.2.2 Risk constrained project portfolio selection

Project portfolio risk management has been studied extensively across various fields and contexts, also covering the aspect of optimal project selection (Micán et al., 2020). Risk constraints are included in portfolio selection problems for multiple reasons. First, regulation may impose limits on the allowed amount of risk or acceptable risk budgets (Baule, 2014). Second, the realization of risky outcomes could inflict significant harm to the underlying organization so that a risk-informed decision maker must enforce a risk constraint. Existing studies on project selection problems have considered various risk measures such as the variance, semi-variance, Value-at-Risk (VaR), Conditional-Value-at-Risk (CVaR), and Gini coefficient (Huang et al., 2014; Hong et al., 2023; Tofighian et al., 2018; Dixit and Tiwari, 2020; Marcondes, 2019). Next, we briefly discuss a few closely related studies on risk constraints in project selection.

Gemici-Ozkan et al. (2010) present a multi-stage decision analysis framework with an abandonment option for selecting a R&D portfolio for a semiconductor manufacturer. They start by quantifying experts' market foresight and qualitative assessments to build a scenario space from which they generate scenarios by sampling. As their model contains multiple decision stages, they use conditional sampling to generate a scenario-tree. This type of scenario generation is also known as nested sampling and is used when modelling multi-stage decisions or random events correlated through time. The authors conduct the multi-stage stochastic optimization routine, in which the expected income is maximized over the generated scenarios. They use a modified Gini coefficient to constrain the total portfolio risk which combines the individual projects' risk measures with a linear function instead of a non-linear one. The authors justify this choice by stating that modeling the true Gini coefficient on the whole portfolio level would have led to a too cumbersome model. The Gini coefficient is a measure of dispersion originally considered as a measure of income equality, but it is also popular risk-measure in different contexts (Yitzhaki and Schechtman, 2013).

Hall et al. (2015) propose a framework for project portfolio selection in which the underlying project return distributions are not explicitly known, although some descriptive statistics of the distributions are available. They justify this choice of modelling uncertainty by remarking that projects are unique and non-recurring. The authors propose a new performance measure called the entropic underperformance riskiness index (URI). They define it as the reciprocal of the Arrow-Pratt absolute risk aversion coefficient for a decision maker who is indifferent between uncertain portfolio returns and the certainty equivalent of these returns given that this certainty equivalent is above a preset threshold. They proceed to minimize this performance measure, which coincides with minimizing risk while achieving the target certainty equivalent.

Tofighian et al. (2018) develop a multi-period portfolio optimization model, in which a deterministic budget is allocated in each time period to a selected number of projects with the possibility of investing left over budget to a risk-free asset. They model the revenues of the projects as normally distributed random variables and constrain the portfolio risk with VaR constraint. The model includes various project interdependencies. For example, when the resources required by two projects are similar, these resources can be shared between the projects to reduce the combined project costs. The authors use a heuristic algorithm to solve the optimization model, as they recognise that no analytical solutions are available due to the complexity of the problem.

Kettunen and Salo (2017) study how uncertainty in project value estimates affects the estimation of risk levels in project portfolio selection. They make connections with the optimizer's curse phenomenon, which states that in optimization problems with uncertainty, the decision maker tends to choose alternatives for which the values are overestimated, leading to post-decision disappointment (see e.g., Smith and Winkler, 2006). They authors proceed to show that uncertain project values can lead to biased Value-at-Risk estimates. Moreover, the inclusion of a risk constraint to the selection problem can lead to an even larger bias, which grows when the risk constraint is made stricter. Nevertheless, they illustrate that the use of a Bayesian calibration approach is effective in removing the bias, and that this calibration can be conducted with Monte Carlo simulation when closed-form posterior distributions are not available.

3 An optimization model for project selection

3.1 Model description

The optimization model presented in this section attempts to capture the key features of the funding mechanism employed by typical public research funding programs. In each time period of the multi-period framework, a fixed budget is available for the funding program to fund research projects. The length of the time periods describes the frequency of decision making and could range anywhere from a few months to several years. In the beginning of a funding period, a set of research project proposals are introduced. These projects proposals can be of different sizes, i.e., some projects need more resources than others. However, a larger, more resource demanding project will also produce more value when completed. As the projects are uncertain, funding large projects can introduce significant concentration risk and can thus be viewed as riskier. On the other hand, funding a few well chosen large projects can also yield great results.

The research projects are granted funding according to the projects' estimated future values. We assume that the future values of the projects are independent of each other. These future values could refer to the possible benefits that the completion of the projects may yield, whether monetary or purely intellectual in nature. Often the two of these are strongly correlated with each other, although it can take a considerable amount of time before scientific findings give rise to useful applications. The estimated future values could be obtained for example by using quantitative technology forecasting techniques (Walk, 2012).

The projects' future values are realized after project completion. We assume that the completion of a project requires funding for its full planned duration. The fact that many research projects fail to produce the outcome that was expected is taken into account in the projects' future values. The decision maker is allowed a one-time option to conduct interim evaluations to a selected subset of on-going projects. This re-evaluation of projects comes at a known cost, and it will produce more accurate estimate of the projects' future values. However, the re-evaluation of projects will not increase the projects' values or affect them in any other way. Using the information from the re-evaluations, the decision maker has the option to abandon some of the projects which will free resources to other projects. Even if a project is abandoned before completion, it may still produce some salvage value, for instance, in the form of an early prototype (Roberts and Weitzman, 1981).

Breakthrough technologies can be seen to stem from excellent research projects with exceptionally large future values. We define excellent projects similarly to that in Öquist and Benner (2012), who considered the top 10% of most cited scientific papers as breakthrough research. More formally, a project is considered excellent if its value exceeds a certain threshold on the project's prior value distribution as illustrated in Figure 1 (a). The probability of this event is referred to as the level of excellence.

However, the value of breakthrough technologies emerging from excellent research projects is not assumed to be fully described by the projects' future values. Instead, we assume that the long-term value of breakthrough technologies is higher than any value attainable by the future value distribution. Therefore, we seek to maximize the number of completed excellent projects in the long-term.

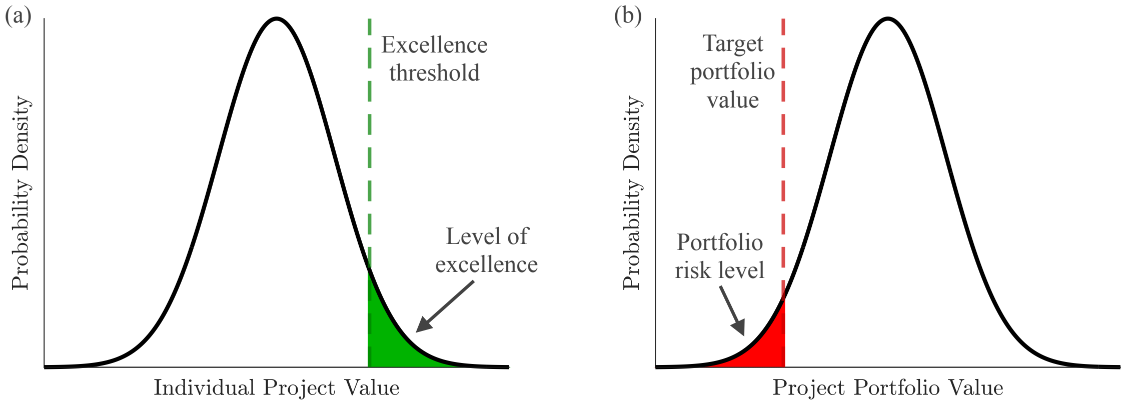

Figure 1: Illustration of (a) the level of excellence and (b) portfolio risk level.

To model different project sizes, we assume that the larger the project the more likely it is to result in a breakthrough technology. This reflects the assumption that projects which require more resources are more ambitious and likely oriented towards less explored research areas, where a significant breakthrough is more probable. Furthermore, to reflect the finding that breakthrough research often requires sustained long-term funding for success (Heinze, 2008), we assume that breakthrough technologies emerge only from completed excellent projects. Thus, we do not consider any salvage value.

We study optimal funding policies for maximizing the number of completed excellent projects in the long-term while also producing satisfactory value in short-term. This value is based on a period-wise target portfolio value which could be set by the funding agency or by an external financier. This target portfolio value should be achieved with high enough probability called the portfolio risk level (see Fig. 1b). In other words, we include a short-term risk constraint by applying the Value-at-Risk framework to the total portfolio value (see e.g., Best, 2000). Furthermore, as we aim to find insight on optimal long-term funding policies for funding risky research, we do not seek to find separate optimal funding policies for each funding period. Instead, we aim to find the optimal stable funding policy, which performs the best on a long time horizon. Such a funding policy is inspired by typical rigid decision-making frameworks that public organizations have. To keep the model sufficiently simple and the computational demands reasonable, we assume that all the different funding periods are identical and do not change in time and that the projects have all equal duration.

3.2 Stable funding policy

In the beginning of each funding period t the decision maker receives a fixed number of proposals for new projects denoted by the set $\mathcal{N}_t$. The deterministic period-wise resource consumption of project $i \in \mathcal{N}_t$ is denoted by c_i and it is dependent on the size of that project. However, the distribution of the projects' period-wise resource consumption is assumed to be stationary and not depend on the funding period t. Thus, the resource commitment needed to fund all projects for one period would be a constant $c_n = \sum_{i \in \mathcal{N}_t} c_i$.

In each funding period, the decision maker chooses to launch a subset $\mathcal{L}_t \subseteq \mathcal{N}_t$ of projects. The number of launched projects $|\mathcal{L}_t|$ may depend on the funding period. The launched projects are selected so that the sum of the period-wise resource consumption of these projects is a fixed amount $c_\ell = \sum_{i \in \mathcal{L}_t} c_i$. This can be considered as the budget for launching new projects. All projects have the same duration d. Therefore, a project launched in the beginning of funding period t will be completed by the start of period $t + d$.

A subset $\mathcal{E}_t \subseteq \mathcal{L}_t$ of the launched projects are funded only until the re-evaluation period $q < d$. The number of re-evaluated projects is

$$|\mathcal{E}_t| = \min(\frac{c_\varepsilon}{e_c}, |\mathcal{L}_t|),$$

where c_ε is the budget allocated for re-evaluation of projects and e_c is the fixed evaluation cost of one project. The rest $\mathcal{L}_t \backslash \mathcal{E}_t$ projects will be granted funding for the whole d periods. Thus, if a project is launched conditionally (i.e., re-evaluated after q periods) it will get funding until period $t + q < t + d$, after which it will proceed to the re-evaluation process. This process is considered to take place within a short time window between the end of period $t + q$ and beginning of period $t + q + 1$.

From the set of re-evaluated projects, the decision maker decides which $\mathcal{A}_t \subseteq \mathcal{E}_t$ projects are abandoned. The size of this set is denoted by $|\mathcal{A}_t|$. The remaining $\mathcal{E}_t \backslash \mathcal{A}_t$ projects are granted funding for the rest $d - q$ periods. The choice to abandon projects is done so that the abandoned projects' total resource consumption $c_a = \sum_{i \in \mathcal{A}_t} c_i$ is constant through different periods.

The total number of completed projects is the sum of projects which are granted full funding and projects which are granted funding after the re-evaluation, or $(|\mathcal{L}_t| - |\mathcal{E}_t|) + (|\mathcal{E}_t| - |\mathcal{A}_t|) = |\mathcal{L}_t| - |\mathcal{A}_t|$. Thus, the total resource commitment for the projects launched in the beginning of funding period t is the sum of (i) resource consumption for $|\mathcal{L}_t|$ projects for q periods, (ii) allocated re-evaluation cost for ε evaluations, and (iii) the resource consumption of $|\mathcal{L}_t| - |\mathcal{A}_t|$ projects for $d - q$ periods. This can be also formulated as $qc_\ell + c_\varepsilon + (d - q)(c_\ell - c_a) = dc_\ell - (d - q)c_a + c_\varepsilon$.

The process of deciding which projects to launch, re-evaluate, and abandon is illustrated in Figure 2. This decision, the stable funding policy, can be summarized with the four decision variables $(c_\ell, c_\varepsilon, c_a, q)$. These decision variables are the same

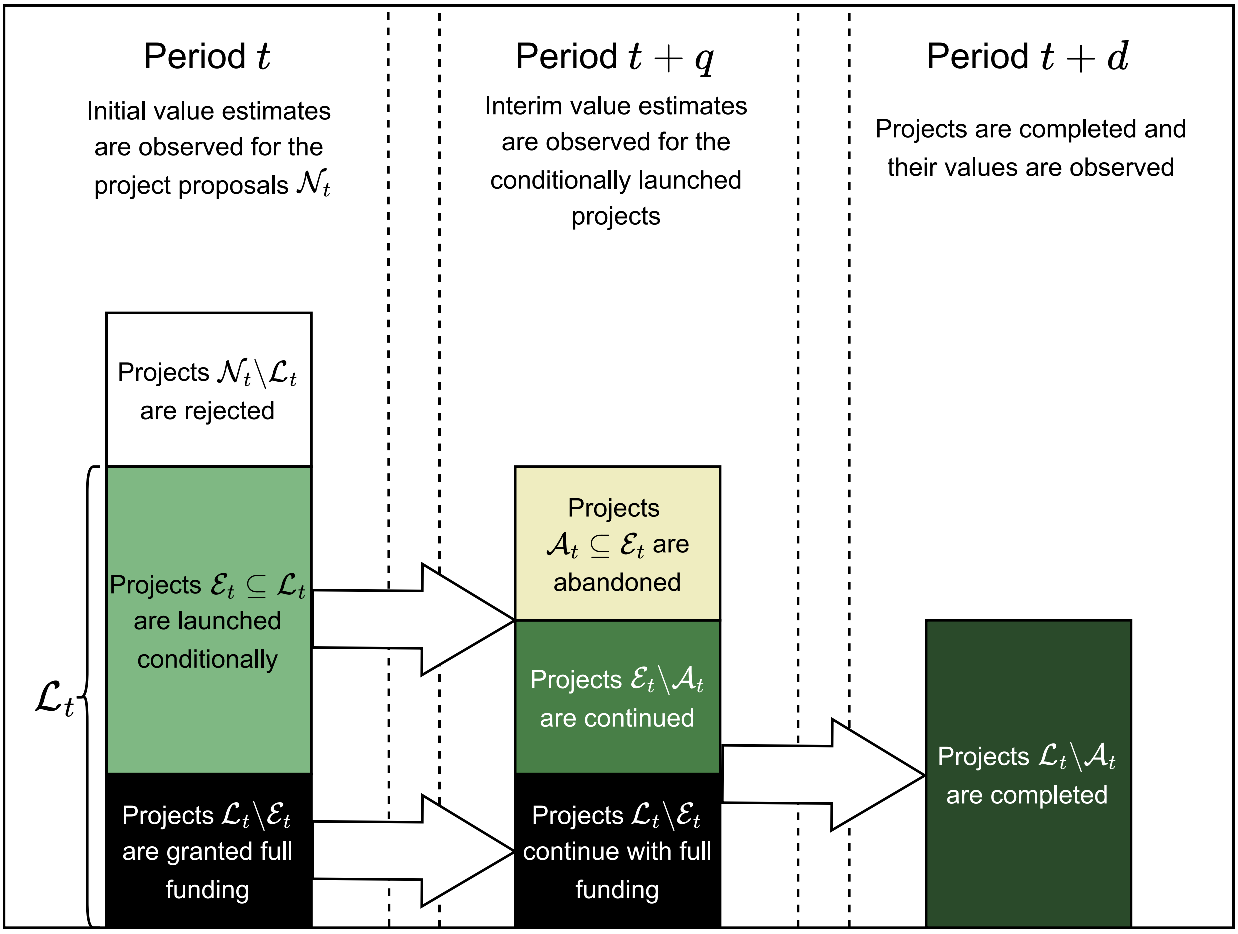

Figure 2: Illustration of the selection and management of projects proposed during funding period t.

for all periods $t \in \mathbb{N}$. In other words, the funding policy is constant over all funding periods t. This means that the period-wise resource commitment is the sum of (i) the resource consumption of projects launched but not yet re-evaluated, (ii) the resource consumption of the on-going re-evaluated projects, and (iii) the evaluation costs allocated for beginning of the period. Thus, the amount of resources needed by the funding policy in period t is

$$C(t) = \begin{cases} tc_\ell & \text{if} \quad 1 \leq t \leq q \\ qc_\ell + (t-q)(c_\ell - c_a) + c_\varepsilon & \text{if} \quad q < t < d \\ qc_\ell + (d-q)(c_\ell - c_a) + c_\varepsilon & \text{if} \quad d \leq t. \end{cases}$$

Note that after the initial project build up lasting d periods, the resource consumption per period will be constant. The stable funding policy is feasible if (i) the resource commitment in each period is withing the budget B, (ii) the resource consumption of abandoned projects does not exceed the resource consumption of launched projects, which in turn does not exceed the resource consumption of all projects, (iii) the re-evaluation budget is a multiple of the evaluation cost of one project with an multiplier less or equal to the number of all projects, and (iv) projects are evaluated at the earliest one period after the launch and at the latest one period before the end. Therefore, the set of all feasible funding policies can be defined as

$$\begin{aligned}\mathcal{P}_F = \big\{(c_\ell, c_\varepsilon, c_a, q) \in \mathbb{N}^4 \mid\ & dc_\ell - (d-q)c_a + c_\varepsilon \leq B,\\ & c_a \leq c_\ell \leq c_n,\\ & c_\varepsilon = ne_c,\ \text{where } n \in \mathbb{N},\\ & 1 \leq q < d\big\}.\end{aligned}$$

Furthermore, not all sets $\mathcal{L}_t$, $\mathcal{E}_t$ and $\mathcal{A}_t$ are feasible. It must hold that the resource consumption of launched and abandoned projects is as described by the funding policy, and the re-evaluation costs are less or equal to allocated costs. Thus, the set of launched, evaluated, and abandoned projects must be included in the following feasible sets in each funding period t:

$$\mathcal{P}_{\mathcal{L}} = \left\{\mathcal{L}_t \subseteq \mathcal{N}_t \mid \sum_{i \in \mathcal{L}_t} c_i = c_\ell \quad \forall t \in \mathbb{N}\right\}$$

$$\mathcal{P}_{\mathcal{E}} = \left\{\mathcal{E}_t \subseteq \mathcal{L}_t \mid \sum_{i \in \mathcal{E}_t} e_c \leq c_\varepsilon \quad \forall t \in \mathbb{N}\right\}$$

$$\mathcal{P}_{\mathcal{A}} = \left\{\mathcal{A}_t \subseteq \mathcal{E}_t \mid \sum_{i \in \mathcal{A}_t} c_i = c_a \quad \forall t \in \mathbb{N}\right\}.$$

3.3 Valuation model

The long term strategy for launching, re-evaluating, continuing, and abandoning re-evaluated projects is determined by the stable funding policy $(c_\ell, c_\varepsilon, c_a, q)$. The decision of selecting the individual projects which will be funded, re-evaluated, and continued in each period t will be based on the projects' estimated future values.

In each time period t there are $|\mathcal{N}_t|$ new project proposals with future value realizations denoted by v_i for $i \in \mathcal{N}_t$. We assume that the projects' future values are conditionally independent and identically distributed (i.i.d) random variables given the projects' resource consumptions c_i. In other words, if two projects have the same period-wise resource consumption, they share the same future value distribution. The prior probability density functions of these future values are known and denoted by $f(v_i \mid c_i)$. Thus, the future value of project i is $V_i \sim f(v_i \mid c_i)$.

The salvage portion is denoted by $h(q)$: $\{1 \,..\, d-1\} \to (0, 1)$. Thus, the portion depends on the re-evaluation period q and we assume it to be non-decreasing in q. If a project is abandoned in period q, a salvage value of $h(q)v_i$ will be obtained. This means that the longer a project is funded, the more value it may potentially yield. However, we assume that there is no additional salvage value emerging from the abandoned projects' excellence.

3.3.1 Initial value estimates

The decision maker is not able to observe the future values of a project directly before project completion. Instead, when the project proposals arrive, the decision maker will observe the estimates s_i^0 for $i \in \mathcal{N}_t$. These estimates will be used to select which projects are launched and funded for the whole d periods and which will be launched and funded conditionally for q periods. We assume that these estimates s_i^0 are i.i.d. when conditioned on the projects' future values v_i and the projects' resource consumptions c_i, i.e., $(S_i^0 \mid V_i = v_i, c_i) \sim f(s_i^0 \mid v_i, c_i)$ with known likelihood probability density distributions $f(s_i^0 \mid v_i, c_i)$ for all possible values of v_i and c_i. Thus, the estimate s_i^0 of project i is independent on the future values of all other projects.

In order to mitigate the optimizer's curse and bias in risk estimates demonstrated by (Kettunen and Salo, 2017), we use Bayesian modeling of probabilities to obtain the posterior estimates of the future values $(V_i \mid S_i^0 = s_i^0, c_i)$. According to the Bayes' rule, we have

$$f(v_i \mid s_i^0, c_i) = \frac{f(s_i^0 \mid v_i, c_i) f(v_i \mid c_i)}{f(s_i^0 \mid c_i)} = \frac{f(s_i^0 \mid v_i, c_i) f(v_i \mid c_i)}{\int_{-\infty}^{\infty} f(s_i^0 \mid v_i, c_i) f(v_i \mid c_i) dv_i}. \tag{1}$$

This formula gives the posterior probability density distribution $f(v_i \mid s_i^0, c_i)$ of the future value $(V_i \mid S_i^0 = s_i^0, c_i)$. Then, the posterior estimates, or the expected values of the projects' future values can be obtained by

$$\mathbb{E}(V_i \mid S_i^0 = s_i^0, c_i) = \int_{-\infty}^{\infty} v_i f(v_i \mid s_i^0, c_i) dv_i.$$

Note that as both a project's future value v_i and posterior value estimate s_i^0 are independent of all other projects' future values and value estimates, the posterior estimate of the project is independent of other projects' future values as well as their prior and posterior value estimates.

3.3.2 Interim value estimates

Similarly to the initial value estimates, by re-evaluating the chosen projects after q periods, we obtain the interim value estimates s_i^q for $i \in \mathcal{E}_t$, which represent the most up-to-date assessment of the projects' future values. These estimates are conditionally independent and identically distributed given the future values v_i. Thus, $(S_i^q \mid V_i = v_i, c_i) \sim f(s_i^q \mid v_i, c_i)$, where the probability density function $f(s_i^q \mid v_i, c_i)$ is known for all values of v_i and c_i. The Bayes' rule can be used to obtain the posterior estimate for the future value $(V_i \mid S^0 = s_i^0, S^q = s_i^q, c_i)$. Note that this time we use both the initial value estimate s_i^0 and s_i^q. Thus, we get

$$f(v_i \mid s_i^0, s_i^q, c_i) = \frac{f(s_i^q \mid v_i, c_i) f(v_i | s_i^0, c_i)}{f(s_i^q \mid c_i)} = \frac{f(s_i^q \mid v_i, c_i) f(v_i \mid s_i^0, c_i)}{\int_{-\infty}^{\infty} f(s_i^q \mid v_i, c_i) f(v_i \mid c_i) dv_i}.$$

Again, we can obtain the expectations of the projects' future values by

$$\mathbb{E}(V_i \mid S_i^0 = s_i^0, S_i^q = s_i^q, c_i) = \int_{-\infty}^{\infty} v_i f(v_i \mid s_i^0, s_i^q, c_i) dv_i.$$

3.3.3 Excellent projects

In alignment with our previous definition of excellent projects, we define them as follows. Project i is an excellent project given that its future value V_i is greater than or equal to the excellence threshold v_i^ξ with a predetermined level of excellence ξ. The threshold is defined in a way that $\mathbb{P}(V_i \geq v_i^\xi) = \int_{v_i^\xi}^{\infty} f(v_i|c_i) dv_i = p(c_i)\xi$, where $p(c_i)$ is an increasing function representing the assumption that projects requiring more resources are more likely to be excellent when completed. Thus, given a project's cost, the parameter ξ defines the desired percentile rank, which determines together with the prior future value distribution whether a given project is excellent or not. The probability that a project is excellent can be calculated as

$$\mathbb{P}(V_i \geq v_i^\xi \mid S_i^0 = s_i^0, c_i) = \int_{v_i^\xi}^{\infty} f(v_i \mid s_i^0, c_i) dv_i$$

during the projects' initial selection, and as

$$\mathbb{P}(V_i \geq v_i^\xi \mid S_i^0 = s_i^0, S_i^q = s_i^q, c_i) = \int_{v_i^\xi}^{\infty} f(v_i \mid s_i^0, s_i^q, c_i) dv_i$$

after the projects have been re-evaluated.

3.4 Optimal funding policy

Next, we formulate the optimization problem by constructing the objective function and the risk constraint. We first establish the connection between maximizing the number of breakthrough technologies in the long-term and the corresponding period-wise expected value.

Proposition 1. *Let $\mathcal{C}_t^{\mathcal{P}}$ be the number of the completed excellent projects launched in the beginning of funding period $t \in \mathbb{N}$ as a function of the funding policy $\mathcal{P} \in \mathcal{P}_F$. Assume that $\mathbb{E}(\mathcal{C}_t^{\mathcal{P}} \mid t \in \mathbb{N})$ is finite with all funding policies. Maximizing the total number of completed excellent projects over a long time horizon with respect to the feasible funding policies $\mathcal{P}_F$ is equivalent to maximizing the expected number of excellent projects $\mathbb{E}(\mathcal{C}_t^{\mathcal{P}} \mid t \in \mathbb{N})$ during each period.*

Proof. We have assumed that the defining features (i.e., the number of project proposals, projects' resource consumptions, projects' future value distributions and observation distributions) of all funding periods are independent and identical over time given a stable funding policy. Therefore, the number of completed excellent projects $(\mathcal{C}_t^{\mathcal{P}} \mid t \in \mathbb{N})$ is a sequence of independent and identically distributed random variables. A realization of these values $(c_t^{\mathcal{P}} \mid t \in \mathbb{N})$ are thus independent and identical random samples.

Let $\mu_n^{\mathcal{P}} = \frac{1}{n}\sum_{t=1}^{n} c_t^{\mathcal{P}}$ for $n \in \mathbb{N}$. According to the strong law of large numbers (see e.g., Ross, 2014), it holds with probability 1 that $\lim_{n\to\infty} \mu_n^{\mathcal{P}} = \mathbb{E}(\mathcal{C}_t^{\mathcal{P}})$. Note that this is true for all funding policies $\mathcal{P} \in \mathcal{P}_F$. Thus, the sequence of functions $\mu_n^{\mathcal{P}}$ converge pointwise to the function $\mathbb{E}(\mathcal{C}_t^{\mathcal{P}})$. Furthermore, this also indicates epi-convergence as all the functions $\mu_n^{\mathcal{P}}$ are bounded real valued functions and the domain of these functions is a finite and discrete subset of $\mathbb{R}^4$. Under these conditions, the theorem of convergence in minimization holds and we can show the equivalence as follows. (Rockafellar and Wets, 2009)

$$
\begin{aligned}
&\lim_{n\to\infty} \arg\max_{\mathcal{P}\in\mathcal{P}_F} \{\sum_{t=1}^{n} c_t^{\mathcal{P}}\} \\
&= \lim_{n\to\infty} \arg\max_{\mathcal{P}\in\mathcal{P}_F} \{\mu_n^{\mathcal{P}}\} \\
&= \lim_{n\to\infty} \arg\min_{\mathcal{P}\in\mathcal{P}_F} \{-\mu_n^{\mathcal{P}}\} \\
&= \arg\min_{\mathcal{P}\in\mathcal{P}_F} \{\lim_{n\to\infty} (-\mu_n^{\mathcal{P}})\} \\
&= \arg\min_{\mathcal{P}\in\mathcal{P}_F} \{-\mathbb{E}(\mathcal{C}_t^{\mathcal{P}})\} \\
&= \arg\max_{\mathcal{P}\in\mathcal{P}_F} \{\mathbb{E}(\mathcal{C}_t^{\mathcal{P}})\}.
\end{aligned}
$$

□

Next, let us denote the set of project proposals arriving at a single funding period simply by $\mathcal{N}$. Furthermore, let the sets $\mathcal{L}$, $\mathcal{E}$, and $\mathcal{A}$ be the projects which are launched, evaluated, and abandoned, respectively. We now construct the objective function for the maximization problem used in finding the optimal long-term funding policies $(c_\ell, c_\varepsilon, c_a, q)$ and the optimal index sets $(\mathcal{L}, \mathcal{E}, \mathcal{A})$, the latter of which depend both on the chosen funding policy and the outlook of the projects' future values.

Using Proposition 1, we can formulate the objective function as the expected number of completed excellent projects. Furthermore, since we assumed that the projects are independent in the sense that their future values and estimates are independent, the expected number of completed excellent projects ω in the portfolio is equal to the sum of the probabilities of the individual projects being excellent:

$$\begin{aligned}
\omega &= \mathbb{E}\left(\sum_{i\in\mathcal{L}\setminus\mathcal{A}} \mathbb{1}(V_i \geq v_i^{\xi})\right) \\
&= \sum_{i\in\mathcal{L}\setminus\mathcal{A}} \mathbb{E}\left(\mathbb{1}(V_i \geq v_i^{\xi})\right) \\
&= \sum_{i\in\mathcal{L}\setminus\mathcal{A}} \int_{-\infty}^{\infty} \mathbb{1}(v_i \geq v_i^{\xi}) f(v_i) dv_i \\
&= \sum_{i\in\mathcal{L}\setminus\mathcal{A}} \int_{v_i^{\xi}}^{\infty} f(v_i) dv_i \\
&= \sum_{i\in\mathcal{L}\setminus\mathcal{A}} \mathbb{P}(V_i \geq v_i^{\xi}).
\end{aligned}$$

Note that here we have not yet conditioned the probabilities using the initial or interim observations. Additionally, when maximizing the number of excellent projects, we do not have to consider the salvage value of the abandoned projects in the objective function. Thus, the optimal funding policy can be found by solving the following two-stage stochastic optimization problem, where we first maximize with respect to the funding policy $(c_\ell, c_\varepsilon, c_a, q)$ and second with respect to the index sets $(\mathcal{L}, \mathcal{E}, \mathcal{A})$:

$$\max_{(c_\ell, c_\varepsilon, c_a, q)\in\mathcal{P}_F} \mathbb{E}_{S_i^0}\left[\max_{\substack{\mathcal{L}\subseteq\mathcal{P}_\mathcal{L} \\ \mathcal{E}\subseteq\mathcal{P}_\mathcal{E}}} \left\{\sum_{i\in\mathcal{L}\setminus\mathcal{E}} \mathbb{P}(V_i \geq v_i^{\xi} \mid S_i^0) + \mathbb{E}_{S_i^q}\left[\max_{\mathcal{A}\subseteq\mathcal{P}_\mathcal{A}} \left\{\sum_{i\in\mathcal{E}\setminus\mathcal{A}} \mathbb{P}(V_i \geq v_i^{\xi} \mid S_i^0, S_i^q)\right\}\right]\right\}\right]. \tag{2}$$

The latter stage can also be seen as a two-stage process. First, the decision is made on the projects $\mathcal{L}$ and $\mathcal{E}$ that are launched and evaluated using the initial value estimates s_i^0. Second, the decision of abandoned projects $\mathcal{A}$ is made using the information from the interim value estimates s_i^q.

3.5 Portfolio risk

The risk of the portfolio $0 < \alpha < 0.5$ can be defined given a target portfolio value $V_\alpha^{\mathcal{P}}$. The portfolio risk is the largest probability for the event that the portfolio value is less than the target value. Thus, the portfolio risk is

$$\alpha = \mathbb{P}\left(V^{\mathcal{P}} \leq V_\alpha^{\mathcal{P}}\right), \tag{3}$$

where $V^{\mathcal{P}}$ is defined by

$$V^{\mathcal{P}} = \mathbb{E}_{S_i^0}\left[\sum_{i\in\mathcal{L}\setminus\mathcal{E}} (V_i \mid S_i^0) + \mathbb{E}_{S_i^q}\left[\sum_{i\in\mathcal{E}\setminus\mathcal{A}} (V_i \mid S_i^0, S_i^q) + \sum_{i\in\mathcal{A}} h(q)(V_i \mid S_i^0, S_i^q)\right]\right]. \tag{4}$$

Thus, $V^{\mathcal{P}}$ is the random variable depicting the value distribution of the whole portfolio given that a certain funding policy is implemented over all possible value observations s_i^0 and s_i^q. Given that the distribution function $f^{\mathcal{P}}(v)$ of the total value of the portfolio $V^{\mathcal{P}}$ is known, the probability in (3) can be calculated by

$$\int_{-\infty}^{V_\alpha^{\mathcal{P}}} f^{\mathcal{P}}(v)dv.$$

In Table 1 we summarize all essential decision variables, key problem parameters, and distributions related to the optimization framework.

1st stage decision variables	
c_ℓ	Launch budget
c_ε	Evaluation budget
c_a	Abandonment budget
q	Evaluation period
2nd stage decision variables	
$\mathcal{L}$	Set of launched projects
$\mathcal{E}$	Set of evaluated projects
3rd stage decision variables	
$\mathcal{A}$	Set of abandoned projects
Key parameters	
$\mathcal{N}$	Set of all projects
v_i	Projects' future values
c_i	Projects' period-wise funding costs
d	Project duration
e_c	Project evaluation cost
B	Total budget
ξ	Level of excellence
v_i^ξ	Excellence threshold
ω	Expected number of excellent projects
α	Level of risk appetite
$V_\alpha^\mathcal{P}$	Target portfolio value
Distributions	
$f(v_i \mid c_i)$	Future value prior probability density function
$f(s_i^0 \mid v_i, c_i)$	Initial value estimate probability density function
$f(s_i^q \mid v_i, c_i)$	Interim value estimate probability density function
$h(q)$	Salvage portion
$p(c_i)$	Resource adjustment to level of excellence

Table 1: Summary of decision variables, key parameters, and distributions.

4 Computation of optimal funding policies

There are no easily derivable closed-form solutions available for the stochastic discrete optimization problem of the model framework outlined in Chapter 3. Therefore, in this Chapter, we present a case problem which we solve numerically using the sample average approximation method (see e.g., Kleywegt et al., 2002). To this end, we present needed analytical results for the case problem. Moreover, we present optimality conditions to reduce the number of viable combinations for decision variables and to improve solution time.

4.1 Sample average approximation method

The idea behind the sample average approximation method (SAA) is simple. First, by means of Monte Carlo simulation, we generate a sample $(W_j \mid j \in \{1 \,..\, \mathcal{S}\})$ of the random variables W included in the optimization problem, where $\mathcal{S}$ is the sample size. In this case, these random variables consist of the projects' initial and interim value observations s_i^0 and s_i^q. Second, we estimate the objective function of the original problem $\mathbb{E}\left[g(x, W)\right]$ with the sample average function $\sum_{j=1}^{\mathcal{S}} \frac{1}{\mathcal{S}} g(x, W_j)$, where x is the decision variable. Then, we maximize the sample average function with respect to x and obtain an estimate of the optimal objective value.

The optimization problem in (2) is essentially a three-stage decision tree. The decisions are made in three stages, in between which some of the random variables are realized and observed by the decision maker. Since the different branches of the decision tree are independent of each other, we can find a solution for the entire problem by combining optimal sub-solutions for the different branches. Thus, we use nested Monte Carlo sampling with two layers reflecting the branching of the decision tree between the different decision stages, illustrated in Figure 3. The first layer contains the sample of the projects' initial value estimates s_i^0 denoted here by $(W_{j_1} \mid j_1 \in \{1 \,..\, \mathcal{S}_1\})$, where $\mathcal{S}_1$ is the sample size of the first layer. In practise, we first generate a sample of the projects' future values v_i using the probability density function $f(v_i \mid c_i)$, after which we use the initial value estimate probability density function $f(s_i^0 \mid v_i, c_i)$ to generate the first layer.

The second layer $(W_{j_1,j_2} \mid j_1 \in \{1 \,..\, \mathcal{S}_1\}, j_2 \in \{1 \,..\, \mathcal{S}_2\})$ of the nested sample consists of the projects' interim value estimates s_i^q. The size of the second layer is $\mathcal{S}_1 \times \mathcal{S}_2$, which means that $\mathcal{S}_2$ samples are generated per each layer one sample. When generating the sample of interim value estimates s_i^q, we must consider that the initial value estimates s_i^0 are already observed by the decision maker. In other words, we generate the sample s_i^q conditioned on the observations s_i^0. In practice, we first obtain a sample of the projects' future values v_i given the sampled initial value estimates s_i^0. The probability density function of this distribution $f(v_i \mid s_i^0, c_i)$ was already described in (1). Then we proceed in generating the sample of the interim value estimates s_i^q using the probability density function $f(s_i^q \mid v_i, c_i)$.

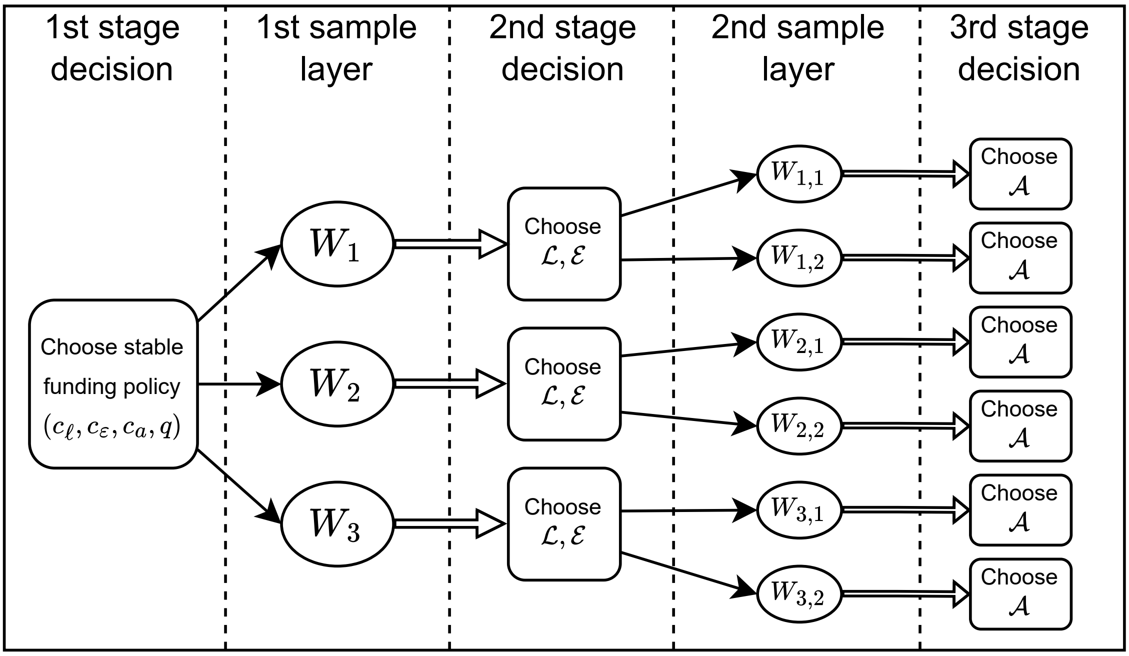

Figure 3: Illustration of the decision tree with sample size of $\mathcal{S}_1 = 3$ on the first layer and $\mathcal{S}_1 \times \mathcal{S}_2 = 3 \times 2 = 6$ on the second layer.

The above method can be used to solve the optimization problem (2) without any constraints. However, we are interested in solving the Pareto front of the optimization problem, which is the set of all non-dominated, Pareto efficient solutions given a constraint based on the portfolio risk level defined in (3). In other words, for each portfolio risk level upper bound $\hat{\alpha}$, we want to obtain the optimal solution ω for the problem (2) given that $\alpha \leq \hat{\alpha}$. However, since this Value-at-Risk based risk measure depends on the full range of outcomes across the whole three-stage decision tree, we cannot form a solution to the original problem by setting a risk constraint for all the sub-problems of the different branches and then combining the solutions.

To overcome this issue, we use the multi-objective optimization framework to form the Pareto front. More specifically, we apply linear scalarization to aggregate the original objective ω and the portfolio risk level α into a single objective function. This is done by summing the variables with weights λ and $1 - \lambda$, where $0 \leq \lambda \leq 1$. Thus, the optimization problem is essentially converted into the form

$$\max \lambda\omega - (1 - \lambda)\alpha,$$

where the second term has a negative sign as we want to minimize the portfolio risk. By searching for the optimal solutions with all values of λ, we form the convex hull of all possible solutions depicting the Pareto front.

Lastly, we discuss the general convergence properties of the sample average approximation method. It can be shown that in quite generic conditions, it holds that

$$\sqrt{\mathcal{S}}(v_s - v^*) \xrightarrow{d} N(0, \sigma^2(x^*)) \text{ when } \mathcal{S} \to \infty,$$

where v^* and x^* are the optimal value and decision variable of the original optimization problem, $\sigma^2(x^*) = \mathrm{Var}[G(x^*, W)]$ is the variance of the objective value with respect to the random element W, and v_s is the optimal value of the approximate optimization problem solved using the sample average approximation method with sample size of $\mathcal{S}$ (Kleywegt et al., 2002). Thus, with a large sample size, it holds roughly that

$$v_s \sim N(v^*, \frac{\sigma^2(x^*)}{\sqrt{\mathcal{S}}}).$$

The accuracy of the solution obtained using SAA-method thus increases with a rate proportional to the square root of the sample size. To balance between feasible solution times and accuracy of the solution, we use a sample size of $\mathcal{S}_1 = 500$ and $\mathcal{S}_2 = 500$ for the first and second layers of the Monte Carlo simulation, respectively.

4.2 Normally distributed project values

Research projects are often seen to produce returns that are heavily skewed, although, there is little evidence on the exact statistical properties (Buzzacchi, 2022). As we already took this skewness of the return distribution into account when modelling excellent projects, we now assume that the projects' future values are normally distributed, i.e., $(V_i \mid c_i) \sim \mathcal{N}(\mu_i, \sigma_i^2)$. Moreover, as we provide general insight on project portfolio selection and not immediate decision recommendations, we believe that normal distributions are the most sensible choice even if in practice the underlying distributions deviate from normality.

We also assume that both the initial and interim estimates $(S_i^0 \mid V_i = v_i, c_i)$ and $(S_i^q \mid S_i^0 = s_i^0, V_i = v_i, c_i)$ are normally distributed. This stems from the idea that these estimates are measurements of the project values with normally distributed error-terms with mean equal to zero. The normality of the error-term can be rationalized by the fact that research project proposals are often reviewed by multiple professionals (Franzoni et al., 2022), and, in the light of the central limit theorem, the averages of those reviews tend to be normal. Formalizing these assumptions, we get $(S_i^0 \mid V_i = v_i, c_i) \sim v_i + \mathcal{N}(0, (\tau_i^0)^2)$ and $(S_i^q \mid V_i = v_i, c_i) \sim \mathcal{N}(v_i, (\tau_i^q)^2)$. Furthermore, to reflect the idea that the interim evaluations are more accurate than the initial evaluations, we set $\tau_i^q = r^{q-1}\tau_i^0$, where $r \in (0, 1)$ is called the uncertainty reduction coefficient. Next, we show that these assumptions lead also to the normality of the posterior value estimates.

Proposition 2. *Let the projects' future values (V_i, c_i) and observed initial value estimates $(S_i^0 \mid V_i = v_i, c_i)$ be normally distributed. Then, the posterior value estimates $(V_i \mid S_i^0 = s_i^0, c_i)$ are also normally distributed. Furthermore, if $(V_i \mid c_i) \sim \mathcal{N}(\mu_i, \sigma_i^2)$*

and $(S_i^0 \mid V_i = v_i, c_i) \sim v_i + \mathcal{N}(0, (\tau_i^0)^2) = \mathcal{N}(v_i, (\tau_i^0)^2)$*, then the distribution of* $(V_i \mid S_i^0 = s_i^0, c_i)$ *has a mean*

$$\mu_i^0 := \frac{s_i^0 \sigma_i^2 + \mu_i (\tau_i^0)^2}{\sigma_i^2 + (\tau_i^0)^2},$$

and standard deviation

$$\sigma_i^0 := \sqrt{\frac{\sigma_i^2 (\tau_i^0)^2}{\sigma_i^2 + (\tau_i^0)^2}}.$$

Similarly, given that $(S_i^q \mid V_i = v_i, c_i) \sim \mathcal{N}(v_i, (\tau_i^q)^2)$*, the distribution of* $(V_i \mid S_i^0 = s_i^0, S_i^q = s_i^q, c_i)$ *has a mean*

$$\mu_i^q := \frac{s_i^q (\sigma_i^0)^2 + \mu_i^0 (\tau_i^q)^2}{(\sigma_i^0)^2 + (\tau_i^q)^2},$$

and standard deviation

$$\sigma_i^q := \sqrt{\frac{(\sigma_i^0)^2 (\tau_i^q)^2}{(\sigma_i^0)^2 + (\tau_i^q)^2}}.$$

Proof. The posterior distribution (1) of the future value of a project given initial value estimate is

$$f(v_i \mid s_i^0, c_i) = \frac{f(s_i^0 \mid v_i, c_i) f(v_i \mid c_i)}{f(s_i^0 \mid c_i)} = \frac{f(s_i^0 \mid v_i, c_i) f(v_i \mid c_i)}{\int_{-\infty}^{\infty} f(s_i^0 \mid v_i, c_i) f(v_i \mid c_i) dv_i}.$$

Thus, as $f(s_i^0 \mid c_i)$ is constant, we have

$$\begin{aligned}
f(v_i \mid s_i^0, c_i) &\propto f(s_i^0 \mid v_i, c_i) f(v_i \mid c_i) \\
&\propto exp\left[-\frac{1}{2}\left(\frac{s_i^0 - v_i}{\tau_i^0}\right)^2\right] exp\left[-\frac{1}{2}\left(\frac{v_i - \mu_i}{\sigma_i}\right)^2\right] \\
&\propto exp\left[-\frac{1}{2}\left(\frac{s_i^0 - v_i}{\tau_i^0}\right)^2 + \left(\frac{v_i - \mu_i}{\sigma_i}\right)^2\right] \\
&\propto exp\left[-\frac{1}{2}\left(\frac{(s_i^0)^2 - 2s_i^0 v_i + v_i^2}{(\tau_i^0)^2} + \frac{v_i^2 - 2v_i\mu_i + \mu_i^2}{\sigma_i^2}\right)\right] \\
&\propto exp\left[-\frac{1}{2}\left(\frac{v_i^2 - 2v_i s_i^0}{(\tau_i^0)^2} + \frac{v_i^2 - 2v_i\mu_i}{\sigma_i^2}\right)\right] \\
&\propto exp\left[-\frac{1}{2}\left(\frac{\sigma_i^2 v_i^2 - 2\sigma_i^2 v_i s_i^0 + (\tau_i^0)^2 v_i^2 - 2(\tau_i^0)^2 v_i \mu_i}{\sigma_i^2(\tau_i^0)^2}\right)\right] \\
&\propto exp\left[-\frac{1}{2}\left(\frac{v_i^2(\sigma_i^2 + (\tau_i^0)^2) - 2v_i(s_i^0\sigma_i^2 + \mu_i(\tau_i^0)^2)}{\sigma_i^2(\tau_i^0)^2}\right)\right] \\
&\propto exp\left[-\frac{1}{2}\left(\frac{v_i^2 - 2v_i\frac{s_i^0\sigma_i^2 + \mu_i(\tau_i^0)^2}{\sigma_i^2 + (\tau_i^0)^2}}{\frac{\sigma_i^2(\tau_i^0)^2}{\sigma_i^2 + (\tau_i^0)^2}}\right)\right] \\
&\propto exp\left[-\frac{1}{2}\left(\frac{v_i - \frac{s_i^0\sigma_i^2 + \mu_i(\tau_i^0)^2}{\sigma_i^2 + (\tau_i^0)^2}}{\frac{\sigma_i\tau_i^0}{\sqrt{\sigma_i^2 + (\tau_i^0)^2}}}\right)^2\right] \\
&\sim N\left(\frac{s_i^0\sigma_i^2 + \mu_i(\tau_i^0)^2}{\sigma_i^2 + (\tau_i^0)^2}, \frac{\sigma_i^2(\tau_i^0)^2}{\sigma_i^2 + (\tau_i^0)^2}\right).
\end{aligned}$$

The second to last step can be done by adding a suitable constant and completing the square (i.e., $v^2 - 2va = v^2 - 2va + a^2 - a^2 = (v-a)^2 - a^2 \propto (v-a)^2$, where a is constant).

The proof for the distribution of $f(v_i|s_i^0, s_i^q, c_i)$ proceeds identically from

$$f(v_i|s_i^0, s_i^q, c_i) \propto f(s_i^q|v_i, c_i) f(v_i|s_i^0, c_i).$$

□

Next, we show how the risk of the project portfolio can be computed given a stable funding policy $\mathcal{P}$ and a target portfolio value $V_\alpha^{\mathcal{P}}$. In Proposition 3, we derive formula for the conditional risk

$$(\alpha \mid S^0 = s^0, S^q = s^q) = \mathbb{P}\left((V^{\mathcal{P}} \mid S^0 = s^0, S^q = s^q) \leq V_\alpha^{\mathcal{P}}\right),$$

which is conditioned on realized posterior value estimates. Using this result, we can simply compute the unconditioned risk α of the portfolio by

$$\alpha = \mathbb{E}_{S^0, S^q}\left[\alpha \mid S^0, S^q\right].$$

Proposition 3. *Let the future value posterior estimates be independent and normally distributed, i.e., let* $(V_i \mid S_i^0 = s_i^0, c_i) \sim \mathcal{N}(\mu_i^0, (\sigma_i^0)^2)$ *for* $i \in \mathcal{L} \setminus \mathcal{E}$ *and* $(V_i \mid S_i^0 = s_i^0, S_i^q = s_i^q, c_i) \sim \mathcal{N}(\mu_i^q, (\sigma_i^q)^2)$ *for* $i \in \mathcal{E} \setminus \mathcal{A}$ *as in Proposition 2. Now, given a stable funding policy* $\mathcal{P}$*, an target portfolio value* $V_\alpha^{\mathcal{P}}$*, and estimates* s_i^0 *for* $i \in \mathcal{L} \setminus \mathcal{E}$ *and* s_i^q *for* $i \in \mathcal{E} \setminus \mathcal{A}$*, it holds that*

$$(\alpha \mid S^0 = s^0, S^q = s^q) = \Phi(\frac{V_\alpha^{\mathcal{P}} - \mu_{total}}{\sigma_{total}}),$$

where

$$\mu_{total} = \sum_{i \in \mathcal{L} \setminus \mathcal{E}} \mu_i^0 + \sum_{i \in \mathcal{E} \setminus \mathcal{A}} \mu_i^q + \sum_{i \in \mathcal{A}} h_i(q) \mu_i^q,$$
$$\sigma_{total} = \sum_{i \in \mathcal{L} \setminus \mathcal{E}} (\sigma_i^0)^2 + \sum_{i \in \mathcal{E} \setminus \mathcal{A}} (\sigma_i^q)^2 + \sum_{i \in \mathcal{A}} (h_i(q) \sigma_i^q)^2$$

and $\Phi(\cdot)$ *is the cumulative distribution function of the standard normal distribution.*

Proof. A well-known result in the field of probability and statistics is that the sum of independent and normally distributed random variables is normally distributed with a mean and variance equal to the sum of means and variances of the original variables, respectively (see e.g., Ross, 2014). Stated in other words, let $X_i \sim \mathcal{N}(a_i, b_i^2)$ for $i \in \mathcal{I}$ be independent and normally distributed and let $X = \sum_{i \in \mathcal{I}} X_i$. Then it holds that $X \sim \mathcal{N}(\sum_{i \in \mathcal{I}} a_i, \sum_{i \in \mathcal{I}} b_i^2)$.

Let $V_{s^0,s^q}^{\mathcal{P}}$ be the total portfolio value as in (4) given the realized posterior value estimates s^0 and s^q. Formally, let

$$\begin{aligned} V_{s^0,s^q}^{\mathcal{P}} &= (V^{\mathcal{P}} \mid S^0 = s^0, S^q = s^q) \\ &= \sum_{i \in \mathcal{L} \setminus \mathcal{E}} (V_i \mid S_i^0 = s_i^0) \\ &\quad + \sum_{i \in \mathcal{E} \setminus \mathcal{A}} (V_i \mid S_i^0 = s_i^0, S_i^q = s_i^q) \\ &\quad + \sum_{i \in \mathcal{A}} h_i(q) (V_i \mid S_i^0 = s_i^0, S_i^q = s_i^q) \end{aligned}$$

Now, as the set of posterior value estimates

$$\Big\{ (V_i \mid S_i^0 = s_i^0, c_i) \mid i \in (\mathcal{L} \setminus \mathcal{E}) \Big\} \bigcup \Big\{ (V_i \mid S_i^0 = s_i^0, S_i^q = s_i^q, c_i) \mid i \in (\mathcal{E} \setminus \mathcal{A}) \Big\}$$

are all independent of each other, we have

$$V_{s^0,s^q}^{\mathcal{P}} \sim \mathcal{N} \left(\sum_{i \in \mathcal{L} \setminus \mathcal{E}} \mu_i^0 + \sum_{i \in \mathcal{E} \setminus \mathcal{A}} \mu_i^q + \sum_{i \in \mathcal{A}} h_i(q) \mu_i^q, \sum_{i \in \mathcal{L} \setminus \mathcal{E}} (\sigma_i^0)^2 + \sum_{i \in \mathcal{E} \setminus \mathcal{A}} (\sigma_i^q)^2 + \sum_{i \in \mathcal{A}} (h_i(q) \sigma_i^q)^2 \right).$$

Let us denote this normal distribution by $\mathcal{N}\left(\mu_{total}, \sigma_{total}\right)$ and let

$$\varphi = \frac{V^{\mathcal{P}}_{s^0,s^q} - \mu_{total}}{\sigma_{total}}$$

be the normalized value of the portfolio. From this it follows that

$$\begin{aligned}
&(\alpha \mid S^0 = s^0, S^q = s^q) \\
&= \mathbb{P}\left((V^{\mathcal{P}} \mid S^0 = s^0, S^q = s^q) \leq V^{\mathcal{P}}_{\alpha}\right) \\
&= \mathbb{P}(V^{\mathcal{P}}_{s^0,s^q} \leq V^{\mathcal{P}}_{\alpha}) \\
&= \mathbb{P}(\frac{V^{\mathcal{P}}_{s^0,s^q} - \mu_{total}}{\sigma_{total}} \leq \frac{V^{\mathcal{P}}_{\alpha} - \mu_{total}}{\sigma_{total}}) \\
&= \mathbb{P}(\varphi \leq \frac{V^{\mathcal{P}}_{\alpha} - \mu_{total}}{\sigma_{total}}) \\
&= \Phi(\frac{V^{\mathcal{P}}_{\alpha} - \mu_{total}}{\sigma_{total}}),
\end{aligned}$$

which completes the proof.

□

4.3 Projects' resource consumption

In its most general form, the optimization problem defined in (2) contains multiple knapsack problems. For instance, choosing the set $\mathcal{L} \subseteq \mathcal{N}$ such that the objective is maximized and that the capacity constraint $\sum_{i\in\mathcal{L}} c_i = c_\ell$ is satisfied is indeed a knapsack problem as the projects' future values are independent. The subproblems of choosing which projects to re-evaluate and which to abandon are also equivalent knapsack problems. Thus, solving these subproblems with many alternative resource consumption levels can be complicated. To avoid solving these knapsack problems for every instance, we restrict the number of different resource consumption levels c_i to two. Let us denote the number of small projects by n_s and the number of large projects by n_l. Thus, $n_s + n_l = |\mathcal{N}|$. Without any loss of generality, we assume that the one period resource consumption for a small project is $c_s = 1$. The resource consumption of a large project is denoted by c_l.

Even with two different types of projects, there is still a vast set of possibilities for choosing the projects to be launched, evaluated, and abandoned. Next, we present optimality conditions illustrating how this set can be reduced. We start by providing a Proposition on how to launch, evaluate, and abandon projects when they are alike.

Proposition 4. *Let the future values V_i and value estimates S_i^0 and S_i^q be normally distributed as in Proposition 2, and let $\mathcal{P} = (c_\ell, c_\varepsilon, c_a, q) \in \mathcal{P}_F$ be a given stable funding policy. Assume all projects have equal resource consumption, i.e., $c_i = c$*

for all $i \in \mathcal{N}$. Assume also that $\frac{c_\ell}{c}, \frac{c_a}{c} \in \mathbb{N}$. Then, the index sets $\mathcal{L} \in \mathcal{P}_\mathcal{L}$, $\mathcal{E} \in \mathcal{P}_\mathcal{E}$, and $\mathcal{A} \in \mathcal{P}_\mathcal{A}$ which maximize the second-stage optimization problem in (2) can be formulated as

$$\begin{aligned} \mathcal{L} &= \{i \in \mathcal{N} \mid \mu_i^0 \textit{ is among the } \frac{c_\ell}{c} \textit{ largest}\} \\ \mathcal{E} &= \{i \in \mathcal{L} \mid \mu_i^0 \textit{ is among the } \frac{c_\varepsilon}{e_c} \textit{ smallest}\} \\ \mathcal{A} &= \{i \in \mathcal{E} \mid \mu_i^q \textit{ is among the } \frac{c_a}{c} \textit{ smallest}\}. \end{aligned} \tag{5}$$

Furthermore, the index sets maximize the expected value of the portfolio μ_{total} and minimize the standard deviation of the portfolio σ_{total}.

Proof. To start with, we show that the index sets in (5) minimize the standard deviation of the portfolio value σ_{total} defined as

$$\sigma_{total} = \sum_{i \in \mathcal{L} \setminus \mathcal{E}} (\sigma_i^0)^2 + \sum_{i \in \mathcal{E} \setminus \mathcal{A}} (\sigma_i^q)^2 + \sum_{i \in \mathcal{A}} (h_i(q)\sigma_i^q)^2.$$

Note that as all projects are alike it holds that $\sigma_i^0 = \sigma_j^0$ and $\sigma_i^q = \sigma_j^q$ for all $i, j \in \mathcal{N}$. It also holds always that $\sigma_i^0 > \sigma_i^q$. This can be shown directly from the definition of σ_i^q (see Prop. 2):

$$\begin{aligned} &\sigma_i^0 > \sigma_i^q \\ \implies &\sigma_i^0 > \sqrt{\frac{(\sigma_i^0)^2(\tau_i^q)^2}{(\sigma_i^0)^2 + (\tau_i^q)^2}} \\ \implies &(\sigma_i^0)^2 > \frac{(\sigma_i^0)^2(\tau_i^q)^2}{(\sigma_i^0)^2 + (\tau_i^q)^2} \\ \implies &(\sigma_i^0)^2 + (\tau_i^q)^2 > (\tau_i^q)^2 \\ \implies &(\sigma_i^0)^2 > 0. \end{aligned}$$

The last step would be false only if $\sigma_i^0 = 0$. This case can be ignored as it would mean that there is no uncertainty in the initial value estimates. Now, due to the feasibility restriction posed by sets $\mathcal{P}_\mathcal{L}$ and $\mathcal{P}_\mathcal{A}$, the number of launched $|\mathcal{L}|$ and abandoned $|\mathcal{A}|$ projects are constants. This means that

$$\begin{aligned} \sigma_{total} &= (|\mathcal{L}| - |\mathcal{E}|)(\sigma_i^0)^2 + (|\mathcal{E}| - |\mathcal{A}|)(\sigma_i^q)^2 + |\mathcal{A}|(h_i(q)\sigma_i^q)^2 \\ &= |\mathcal{L}|(\sigma_i^0)^2 + |\mathcal{E}|((\sigma_i^q)^2 - (\sigma_i^0)^2) + |\mathcal{A}|((h_i(q)\sigma_i^q)^2 - (\sigma_i^q)^2). \end{aligned}$$

The only non-constant term here is $|\mathcal{E}|((\sigma_i^q)^2 - (\sigma_i^0)^2)$. Choosing to evaluate $\frac{c_\varepsilon}{e_c}$ projects maximizes $|\mathcal{E}|$ which in turn minimizes this only non-constant term, and, therefore, also minimizes the standard deviation of the portfolio value σ_{total}.

To show that the index sets in (5) maximize the number of excellent projects (2) and expected portfolio value μ_{total}, we refer to the proof of Proposition 1 by Vilkkumaa et al. (2015) which starts from the same set of assumptions.

□

Next, we use Proposition 4 to construct all possible index sets $\mathcal{L}$, $\mathcal{E}$, and $\mathcal{A}$ among which we can search for the optimal alternatives in the case where there are two different types of projects. We start by dividing the first three decision variables $(c_\ell, c_\varepsilon, c_a)$ into parts. Let us denote $(c_\ell^s, c_\varepsilon^s, c_a^s)$ and $(c_\ell^l, c_\varepsilon^l, c_a^l)$ as the budgets for launching, evaluating, and abandoning small and large projects, respectively. It must hold that

$$\begin{aligned} c_\ell &= c_\ell^s + c_\ell^l \\ c_\varepsilon &= c_\varepsilon^s + c_\varepsilon^l \\ c_a &= c_a^s + c_a^l. \end{aligned} \tag{6}$$

The problem of choosing which projects to launch, evaluate, and abandon is separated into two independent problems. The solutions $\mathcal{L}^s$, $\mathcal{L}^l$, $\mathcal{E}^s$, $\mathcal{E}^l$, $\mathcal{A}^s$ and $\mathcal{A}^l$ for the problems can be obtained using Proposition 4. Let ω^s, ω^l, μ_{total}^s, μ_{total}^l, σ_{total}^s and σ_{total}^l be the corresponding expected number of excellent projects, expected values, and standard deviations of the portfolio values given the optimal subsolutions. We can combine these solutions as

$$\begin{aligned} \mathcal{L} &= \mathcal{L}^s \cup \mathcal{L}^l \\ \mathcal{E} &= \mathcal{E}^s \cup \mathcal{E}^l \\ \mathcal{A} &= \mathcal{A}^s \cup \mathcal{A}^l \\ \omega &= \omega^s + \omega^l \\ \mu_{total} &= \mu_{total}^s + \mu_{total}^l \\ \sigma_{total} &= \sqrt{(\sigma_{total}^s)^2 + (\sigma_{total}^l)^2}. \end{aligned}$$

Given the initial partition of the problem into the two subproblems, this is the solution resulting in both maximal ω and μ_{total}, and minimal σ_{total}. Next, we show that this also minimizes the conditional risk of the portfolio. Here we assume that the target portfolio value is conservative enough so that $V_\alpha^{\mathcal{P}} < \mu_{total}$. Then, it holds that

$$
\begin{aligned}
&\min\left\{\alpha \mid S^0 = s^0, S^q = s^q\right\} \\
=&\min\left\{\Phi(\frac{V_\alpha^{\mathcal{P}} - \mu_{total}}{\sigma_{total}})\right\} \\
=&\min\left\{\frac{V_\alpha^{\mathcal{P}} - \mu_{total}}{\sigma_{total}}\right\} \\
=&\frac{\min\left\{V_\alpha^{\mathcal{P}} - \mu_{total}\right\}}{\min\sigma_{total}} \\
=&\frac{V_\alpha^{\mathcal{P}} - \max\mu_{total}}{\min\sigma_{total}}.
\end{aligned}
$$

Thus, to obtain the optimal index sets $\mathcal{L}$, $\mathcal{E}$, and $\mathcal{A}$ which maximize the expected number of excellent projects and minimize risk, it suffices to compute the optimal solutions for the different ways to divide the budgets for the small and large projects. For instance, if the number of large projects is $n_l = 5$, there are $\sum_{i=0}^{5}\sum_{j=0}^{i}\sum_{k=0}^{j} 1 = 56$ different ways to choose how many of them is launched (i), evaluated (j), and abandoned (k). These choices fix the budgets $(c_\ell^l, c_\varepsilon^l, c_a^l)$ for the large projects, and, through (6), they also fix the budgets for the small projects. Thus, given the stable funding policy $\mathcal{P}$, there is a maximum of 56 cases that need to be considered.

Lastly, we give a result useful for restricting the number of combinations for first-stage decision variables $\mathcal{P} = (c_\ell, c_\varepsilon, c_a, q)$.

Proposition 5. *Let the stable funding policy $\mathcal{P} = (c_\ell, c_\varepsilon, c_a, q)$ be optimal and assume that the number of available evaluations $\frac{c_\varepsilon}{e_c}$ is less than the launching budget c_ℓ. Then, the funding policy $\mathcal{P} = (c_\ell, c_\varepsilon + 1, c_a, q)$ is infeasible.*

Here, we refer to the rigorous proof provided by Vilkkumaa et al. (2015). The idea behind the Proposition is quite simple. Assume that all other decision variables are fixed. Then, the evaluation budget should be as large as the possible, as the possibility to evaluate more projects cannot be disadvantageous with respect to the objectives of the problem. If the current funding policy already makes it possible to evaluate all projects, then no more budget should be allocated to evaluations but, rather to launch more projects or to abandon fewer projects.

5 Results

In this Chapter, we present the numerical results of the optimization model described in Chapter 3 to gain insight on optimal project selection and management. More specifically, we seek to answers to the following research questions:

(i) How can the decision maker take into account the different projects' sizes and risks when selecting projects?

(ii) How should the decision maker's appetite for risk be reflected in the funding strategy?

(iii) How can the abandonment option be utilized in developing breakthrough technologies under concentration risk and as a risk management tool?

We start by providing solutions the optimization problem (2) in varying conditions. Then, we compare alternative project selection and management strategies with optimal solutions.

5.1 Parameters

Table 2 presents the model parameters and functions which act as a starting point of our analysis. We consider two different levels of resource consumption for the projects as already mentioned in Chapter 4.3. The level of resource consumption is $c_s = 1$ and $c_l = 10$, and the number of project proposals in each period is $n_s = 50$ and $n_l = 5$ for small and large projects, respectively. We set the duration of all projects to $d = 4$. Thus, the resource requirement for funding all projects for the full 4 periods is $4 \times (50 \times 1 + 5 \times 10) = 400$. We set the budget B so that 25% of the projects can be funded until completion assuming no projects are evaluated, i.e., $B = 0.25 \times 400 = 100$. The cost of evaluating one project is $e_c = 0.1$.

The means and standard deviations of the projects' future values and the estimation errors are assumed to scale linearly with respect to the level of resource consumption. Thus, the projects' future values are normally distributed with means $\mu_s = 20$ and $\mu_l = 200$, and standard deviations $\sigma_s = 3$ and $\sigma_l = 30$ for small and large projects, respectively. The estimation errors related to the initial value observations are $\tau_s^0 = 7$ and $\tau_l^0 = 70$, and the corresponding errors of the interim observations are $\tau_s^q = 7r^{q-1}$ and $\tau_l^q = 70r^{q-1}$, where the uncertainty reduction coefficient is $r = 0.5$. Thus, the

Parameter	**Value**
c_s	1
c_l	10
n_s	50
n_l	5
B	100
μ_s	20
μ_l	200
σ_s	3
σ_l	30
τ_s^0	7
τ_l^0	70
r	0.5
d	4
e_c	0.1
ξ	0.02
$V_\xi^{\mathcal{P}}$	470
Function	**Value**
$h(q)$	$\left(\frac{q-1}{d}\right)^{1.3}$
$p(c_i)$	c_i

Table 2: Parameters and functions

distributions of the future values and their observations are

$$\begin{aligned}
(V_i \mid c_s) &\sim N(20, 3^2) \\
(V_i \mid c_l) &\sim N(200, 30^2) \\
(S_i^0 \mid v_i, c_s) &\sim N(v_i, 7^2) \\
(S_i^0 \mid v_i, c_l) &\sim N(v_i, 70^2) \\
(S_i^q \mid v_i, c_s) &\sim N(v_i, (7r^{q-1})^2) \\
(S_i^q \mid v_i, c_l) &\sim N(v_i, (70r^{q-1})^2).
\end{aligned}$$

The salvage portion function is set to be slightly convex and is $h(q) = (\frac{q-1}{d})^{1.3}$. This reflects the assumption that projects tend to start slowly, and major progress is made during the projects' final stages. The level of excellence is set at $\xi = 2\%$ and we assume that it increases linearly with resource consumption. Thus, we set $p(c_i) = c_i$ implying that a small project is excellent with a probability of $1 \times 2\% = 2\%$ and a large project is excellent with probability $10 \times 2\% = 20\%$. The target portfolio value is set at $V_\alpha^{\mathcal{P}} = 470$. This is 6% less than $25 \times 20 = 500$, which is the expected portfolio value when randomly selecting 25, or the maximum number of small projects to the portfolio each with an expected value of 20. The optimization model is solved using MATLAB R2021a and a nested Monte Carlo sample size of 500×500.

5.2 Optimal funding policy characteristics

Figure 4 presents four funding policies from the Pareto front for a fixed evaluation period $q = 3$, indicating that the interim evaluations are conducted in the middle of the projects' life cycles. In each panel (a)-(d), we display how many small and large projects are on average launched, evaluated, and abandoned given a stable funding policy $\mathcal{P} = (c_\ell, c_\varepsilon, c_a, q)$. Here, we focus only on the projects which are launched on a single funding period $t = 1$ and do not consider projects launched during later funding periods. In the first period, the decision maker can either reject projects (R), launch projects without possibility to re-evaluate (FF), or launch projects with the option to re-evaluate (CF). In period 3, the decision maker can choose to continue (C) or abandon (A) the re-evaluated projects.

With the first and most risk-averse funding policy (a), the probability that the portfolio value is less than the target portfolio value $V_\alpha^{\mathcal{P}} = 470$ is close to zero. This is indicated by the portfolio risk level $\alpha \approx 0$. The Pareto efficient funding policy $\mathcal{P} = (c_\ell, c_\varepsilon, c_a, q) = (25, 0, 0, 3)$ denotes that the whole budget is spent on launching projects. Furthermore, only small projects are launched. This funding policy can be seen as the most resource conserving policy as no resources are spent on experimenting with projects, some of which would have to be abandoned later. Moreover, this is the most diversified funding policy with respect to the number of completed projects. Both choices lead to the very small portfolio risk level. However, the excepted number of completed excellent projects is also the smallest at $\omega = 0.85$ projects per period.

The funding policy (b) $\mathcal{P} = (25, 0, 0, 3)$ is described by the same first-stage decision variables as funding policy (a). The difference comes from how the budget is allocated between the small and large projects. In funding policy (b), on average 0.8 large project are launched across the 500 simulated instances together with, on average, 17 small projects. The value threshold for choosing whether to launch or reject a large project is roughly $\mathbb{E}[V_i \mid s_i^0] = 210$. Due to this, 2 large projects are launched in 6% of the instances, 1 large project in 66% of the instances, and no large projects in 28% of the instances. The probability that the target portfolio value is not reached is approximately $\alpha = 0.01$ and the excepted number of completed projects is $\omega = 1.01$. This seems to indicate that by having the opportunity to launch large projects when they seem highly promising, the decision maker can obtain notably higher expected value of completed excellent projects while keeping the portfolio risk on a suitable level as many small projects are also funded.

In funding policy (c) $\mathcal{P} = (30, 2, 11, 3)$ almost all projects are funded conditionally, i.e., funding is initially granted until the projects' re-evaluation in the beginning of period 3. On average, the number of launched large projects is 0.9, which is similar as with funding policy (b). However, now roughly one third of the launched projects are abandoned on period 3. Thus, while the number of launched projects is increased by 20%, the number of completed projects is, vice versa, reduced by roughly 20%. This increases the chance to launch and then continue excellent projects, yet

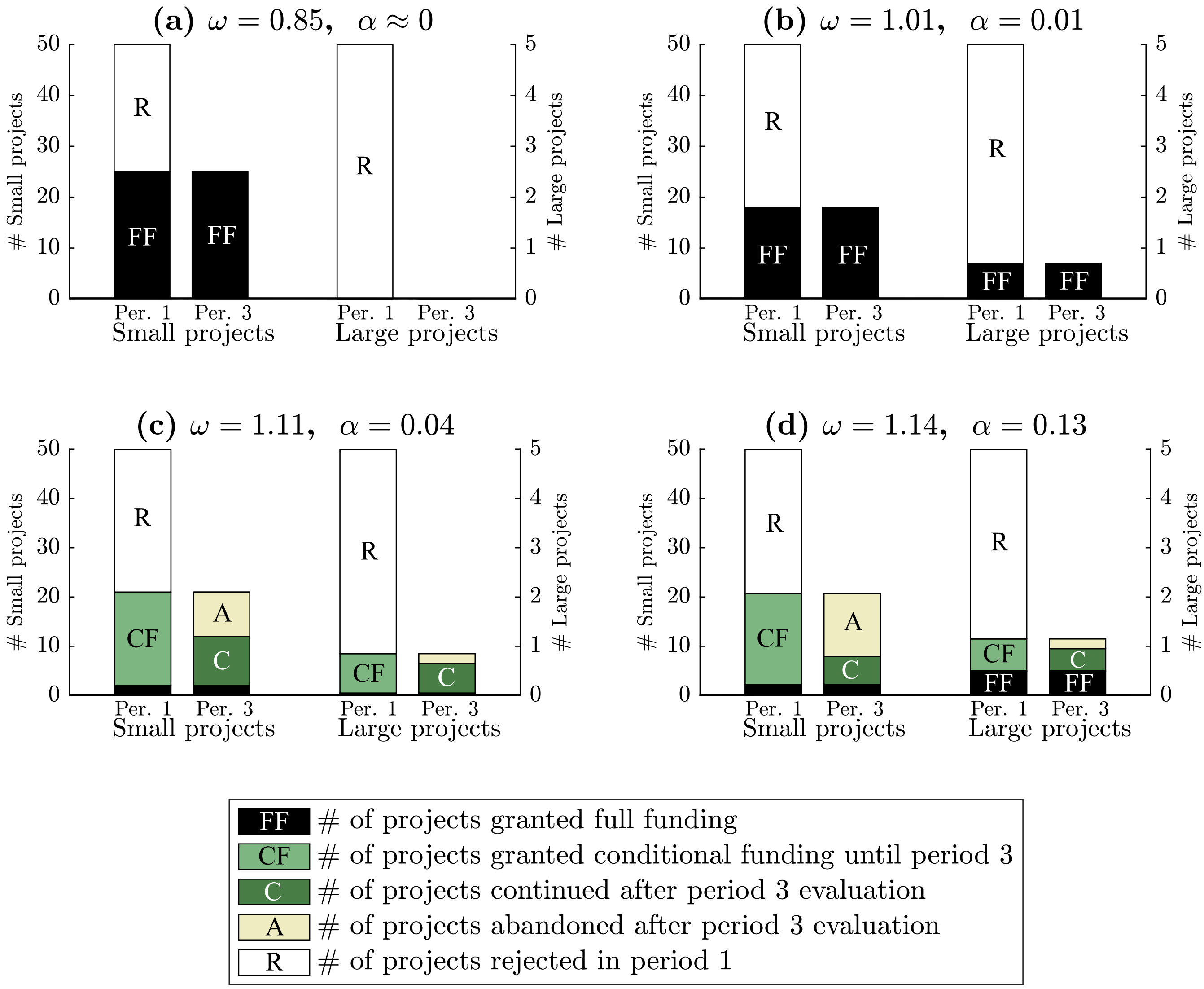

Figure 4: Illustration of the period 1 and period 3 actions of four funding policies from the Pareto front with portfolio risk levels (a) $\alpha \approx 0$, (b) $\alpha = 0.01$, (c) $\alpha = 0.04$, and (d) $\alpha = 0.13$. The size of the nested Monte Carlo simulation is 500×500.

simultaneously reduces the amount of final diversification. Furthermore, while 40% of the launched small projects are abandoned in period 3, only 27% of large projects are abandoned. This indicates that large projects are still launched quite cautiously when compared to small projects. The use of the abandonment option increases the objective value roughly 10% to $\omega = 1.11$ compared to funding policy (b). The portfolio risk level is also increased and is $\alpha = 0.04$.

Figure 5 illustrates the conditions in which different decisions regarding large projects are made under funding policy (c). The first panel (i) represents the period 1 decision, showing that large projects which are granted full funding for the whole 4 periods

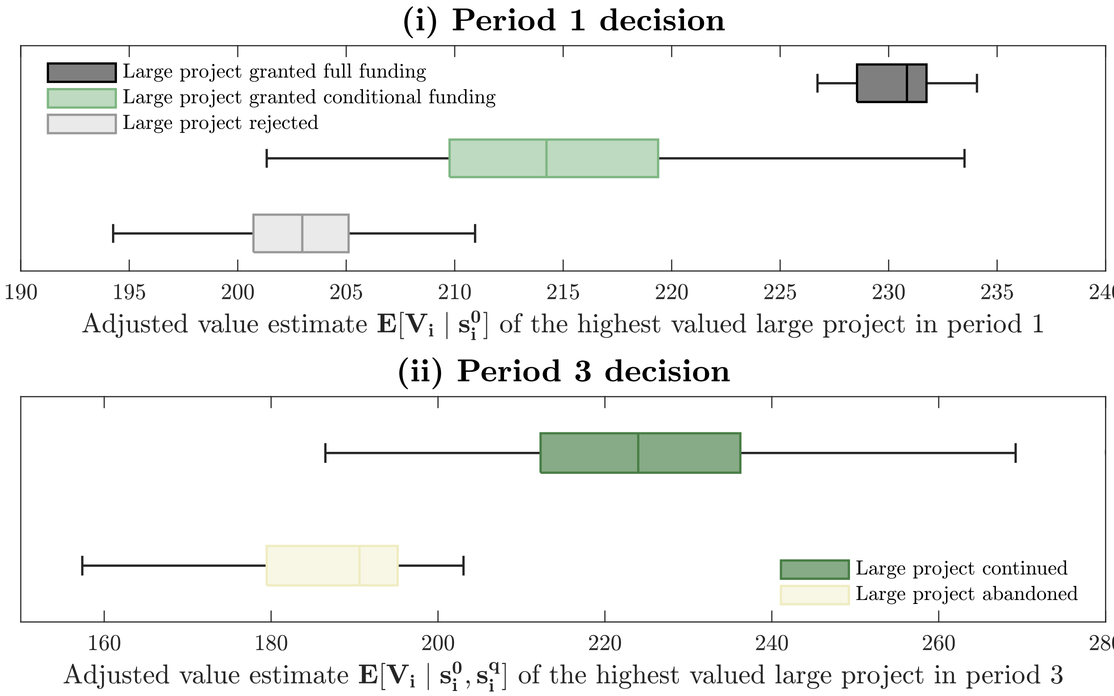

Figure 5: The range of the top 1 ranked large projects' value estimates over 500 simulations separated according to both (i) period 1 and (ii) period 3 decisions according to funding policy (c). A standard boxplot is used, i.e., the line inside each box represents the median value, the edges of the box represent the 0.25 and 0.75 quantiles, the whiskers represent the maximum and minimum non-outlier values defined using the interquartile range. Outliers are left out of the visual for clarity.

are estimated to have very high values roughly around $\mathbb{E}[V_i \mid s_i^0] = 230$. Projects with this large value estimates are quite likely to be excellent, which is why the abandonment option is not used. On the other hand, large projects are often rejected when the value estimates are less than 205.

The second panel (ii) in Figure 5 shows that after re-evaluation, large project are continued when the updated value estimates $\mathbb{E}[V_i \mid s_i^0, s_i^q]$ are approximately over 200. Below this value, large projects are very unlikely to be excellent, which is why the best decision is often to abandon them and use the leftover resources to continue 10 small projects instead. Note that the value estimates of large projects do not alone explain the decisions concerning large projects but the whole portfolio view has to be considered. For instance, if the set of project proposals is poor overall, it can be too risky to launch even quite promising large projects.

The last, most risk-seeking funding policy (d) is $\mathcal{P} = (32, 2, 15, 3)$. The expected number of completed excellent projects with this policy is $\omega = 1.14$, which is only a small improvement compared to funding policy (c). The probability that the target portfolio value is not reached is $\alpha = 0.13$. Now, almost half of all launched projects are abandoned in period 3. Furthermore, somewhat surprisingly, almost half of the large projects are launched unconditionally (i.e., granted full funding). However, when comparing to funding policy (c), we see that the risk-reward ratio of not having the option to re-evaluate large projects is quite poor as this increases the objective only by 0.03 while tripling the probability of not achieving the target portfolio value. Thus, it seems that the main role of the wide use of the abandonment option for large projects in funding policy (c) was reducing the portfolio risk. This is clearly seen in Figure 6 (i), which shows the Pareto front associated with the optimization problem including the four discussed funding policies (a)-(d). The Pareto front is convex, which means that the benefit of switching to a riskier funding policy decreases as the risk of the portfolio increases. While the overall optimal funding policy depends on the risk appetite and preferences of the decision maker, the funding policy (d) can still be considered as non-optimal in most cases.

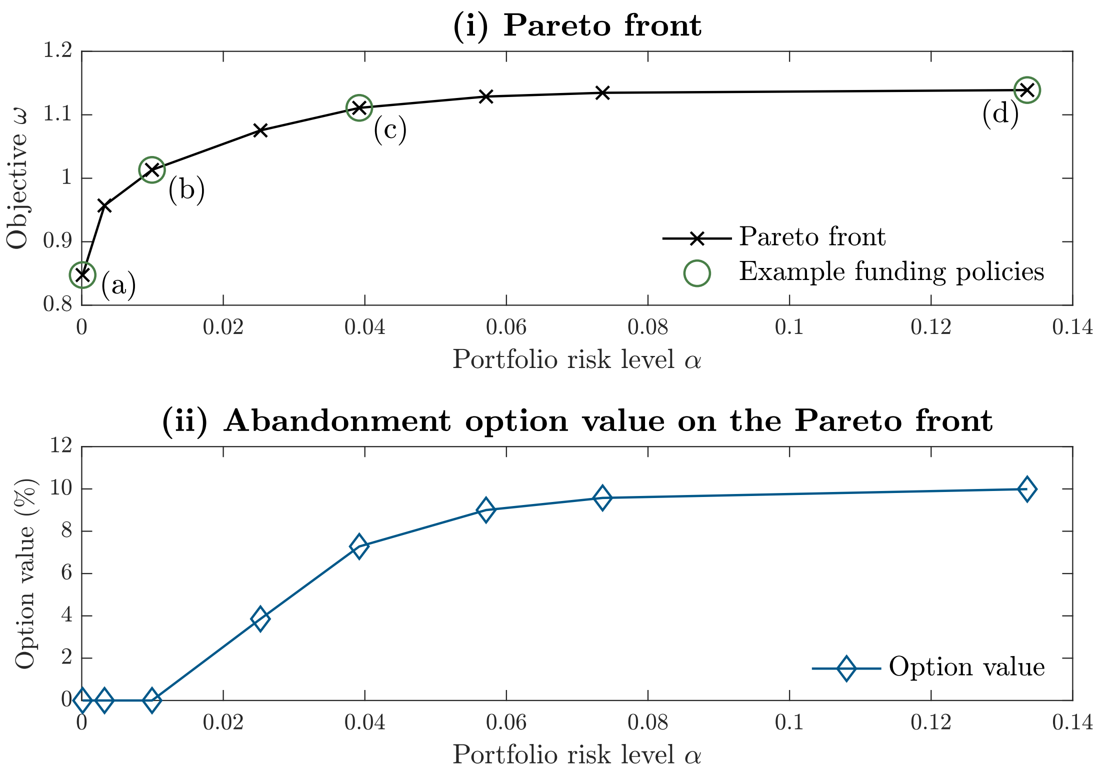

Figure 6: (i) The Pareto front of funding policies including the funding policies (a)-(d) presented in Figure 4 and (ii) the abandonment option value at the Pareto front.

Figure 6 (ii) illustrates the abandonment option value with different levels of portfolio risk. The option value is defined as the relative difference in the objective value ω given that the abandonment option is available when the baseline is the objective value without the option. In the model, the option can be removed simply by setting

$c_\varepsilon = c_a = 0$. Thus, the abandonment option value is

$$\text{Option value} := \frac{\text{Objective value with option - Objective value without option}}{\text{Objective value without option}}.$$

Figure 6 (ii) shows that the option value is an increasing function with respect to the portfolio risk level α. A decision maker with a low level of risk appetite will not benefit at all from the option, whereas a decision maker with higher tolerance for risk can expect up to a 10% increase in the number of completed excellent projects. This is a bit unexpected, as the abandonment option tends to reduce the risk of committing funding for individual projects which perform poorly. However, here the use of the abandonment option may not be flexible enough for this effect to materialize as the same proportion of projects should be abandoned during each period.

5.2.1 Varying budget

Figure 7 presents the effect of budget size B to optimal funding strategy with two levels of portfolio risk of $\alpha = 0.10$ and $\alpha = 0.03$. The target portfolio value $V_\alpha^{\mathcal{P}}$ is varied in proportion to the budget so that it remains roughly 6% below the expected portfolio value when granting full funding to a maximum number of small projects. Panel (i) shows how the portion of budget spent on launching large projects increases with both risk levels as the budget increases. There are a few explanatory factors behind this phenomenon. First, as the budget increases, the decision maker is allowed to launch more projects leading to a lower portfolio risk through improved diversification. The reduction in risk allows the decision maker to steer more of the funding towards large projects. On the other hand, as the decision maker is allowed to launch more projects, the average quality of these projects will decrease as better project are granted funding first, making the selection of alternative large projects more sensible.

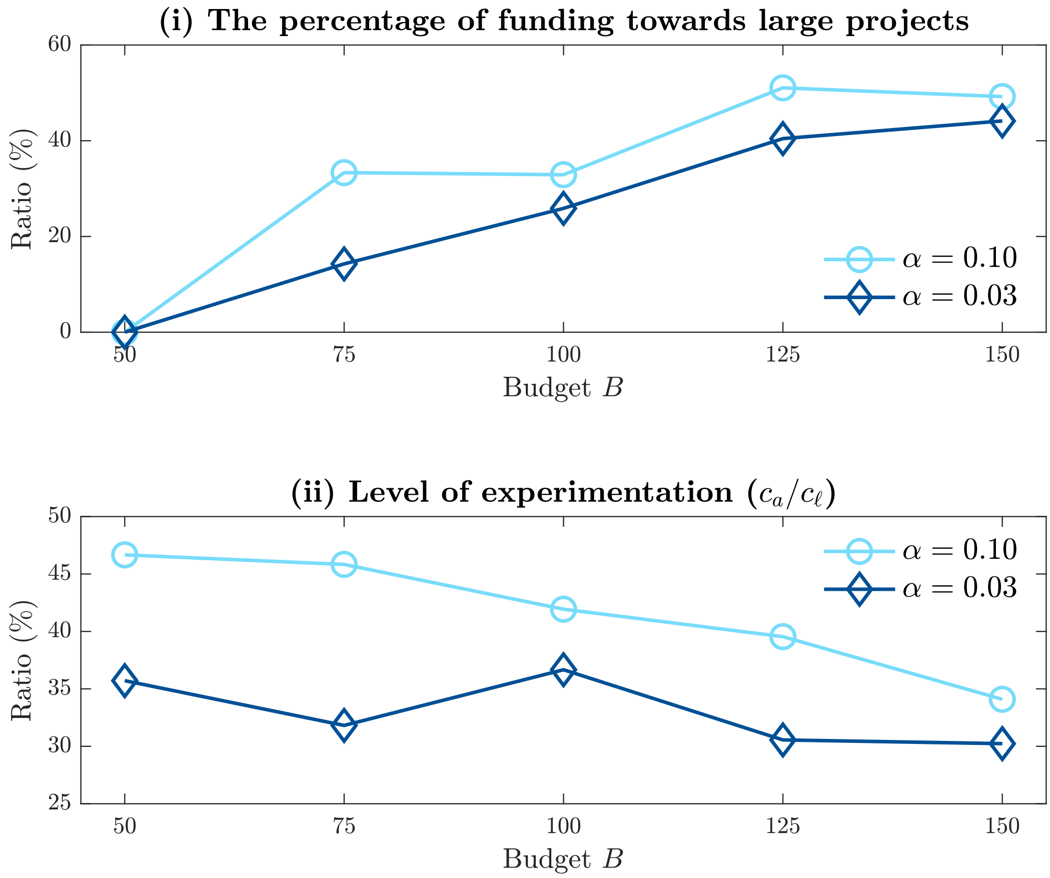

Figure 7: The effect of budget B to (i) the percentage of funding towards large projects, and (ii) the level of experimentation defined as the ratio of the abandonment budget and launch budget.

Figure 7 (ii) shows the effect of budget to the optimal funding policy's level of experimentation, defined as the ratio of the abandonment budget c_a and the launch budget c_ℓ. The larger the level of experimentation, the more projects are both initially launched and abandoned after re-evaluation. After re-evaluation, there is thus a larger selection of projects from which the completed projects can be selected. In general, the level of experimentation decreases as the budget increases. However, there is one data point not following this trend due to numerical model limitations. The trend is more obvious with the higher portfolio risk level $\alpha = 0.10$. This may be counterintuitive as with a larger budget there are more resources to spend on experimenting with a larger set of projects. However, a larger budget decreases the average project quality, making it less sensible to experiment with unpromising projects than to grant full funding to the most promising ones. On the contrary, with a smaller budget more experimentation should be done to increase the likelihood of funding excellent projects.

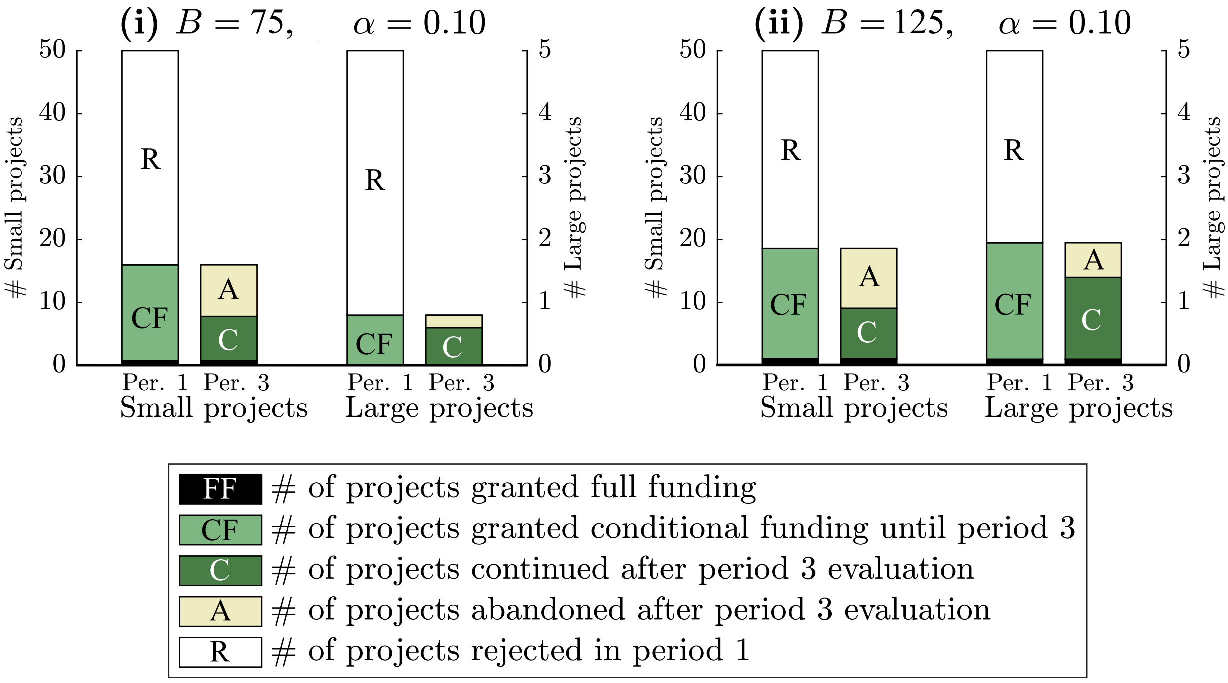

Figure 8: Illustration of two funding policies with budgets (i) $B = 75$ and (ii) $B = 125$. Both funding policies result in a portfolio risk level of $\alpha = 0.10$

Figure 8 illustrates two funding policies with portfolio risk level $\alpha = 0.10$. The first funding policy (i) with budget $B = 75$ is $\mathcal{P} = (24, 2, 11, 3)$, with one large projects being launched in 80% of instances and no large projects in the rest. The funding of large projects is cautious as they are launched conditionally and only 25% of them are abandoned after re-evaluation. The second funding policy (ii) with increased budget $B = 125$ spends the additional budget mostly to launch more large projects. One large project is launched in 13% of instances, two large projects in 80% of instances, and three large projects in 7% of instances. Still roughly 25% of large projects are abandoned. This indicates that the threshold for launching large projects is significantly reduced along with the threshold of continuing them after re-evaluation. This can be done as there is already, on average, 20 small projects funded in each period resulting in sufficient diversification.

5.2.2 Choice of the evaluation period

Next, we free the decision variable q denoting the re-evaluation period. Previously, the re-evaluations were conducted half-way through the projects' life cycles in the beginning of period 3. Here, we set $\tau_s^0 = 10$ and $\tau_l^0 = 100$ giving the following initial and interim value observation distributions:

$$
\begin{aligned}
(S_i^0 \mid v_i, c_s) &\sim N(v_i, 10^2) \\
(S_i^0 \mid v_i, c_l) &\sim N(v_i, 100^2) \\
(S_i^q \mid v_i, c_s) &\sim N(v_i, (10r^{q-1})^2) \\
(S_i^q \mid v_i, c_l) &\sim N(v_i, (100r^{q-1})^2).
\end{aligned}
$$

Figure 9 (i) presents the Pareto front with free re-evaluation period q. The vertical line at $\alpha = 0.06$ represents the threshold where the re-evaluation period changes from period $q = 2$ to period $q = 3$ in the optimal funding policy. When the level of risk appetite is smaller than $\alpha = 0.06$, conditionally launched projects are re-evaluated after one experimentation period. When projects are evaluated early, more of them can be completed as less resources are needed for the experimentation. On the other hand, the interim value estimates become less accurate increasing the uncertainty of their value. The level of experimentation (c_a/c_ℓ) also increases along the Pareto front as illustrated in Figure 9 (ii). Thus, with an increased risk appetite, it is beneficial to both experiment longer and with more projects. With longer experimentation, the decision maker is more likely to detect excellent projects as the interim value estimates are more accurate, yet more resources are wasted on funding projects which are not completed. Figure 9 (ii) illustrates also how the probability of getting a portfolio value less than 470 increases along with the proportion of large projects in the portfolio as diversification worsens.

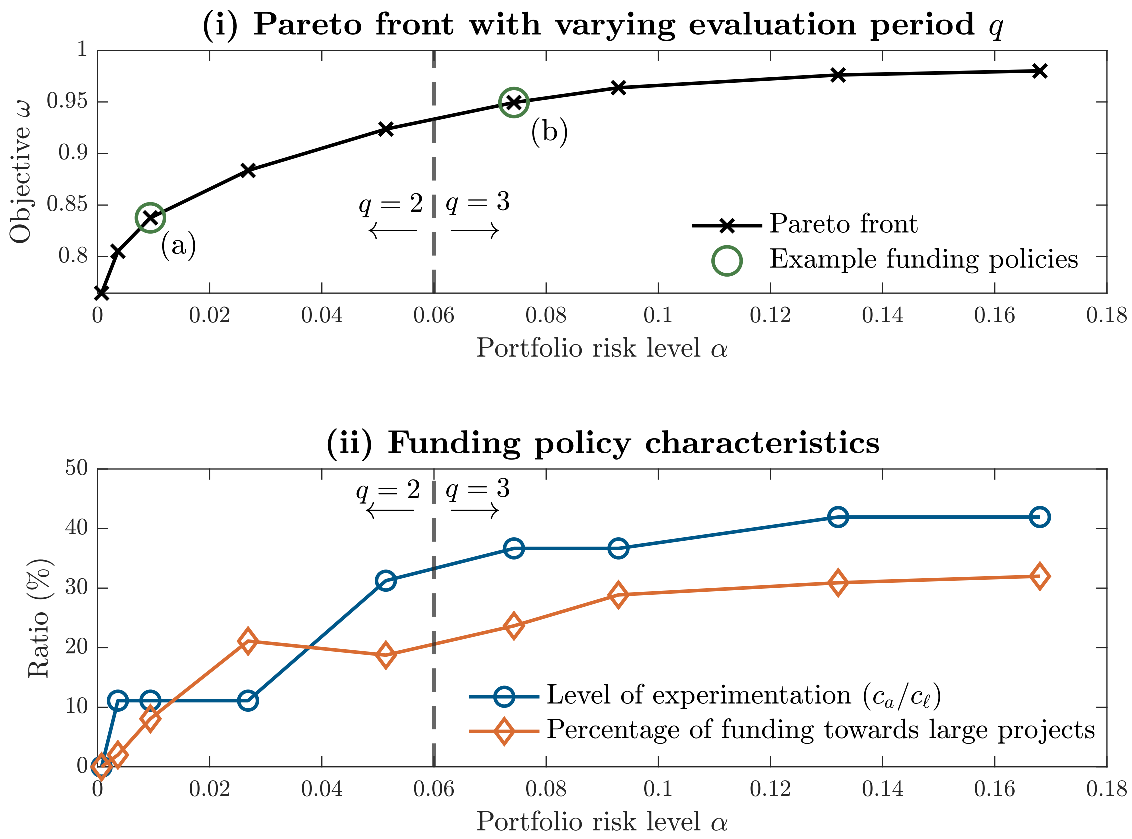

Figure 9: (i) The Pareto front of funding policies with free re-evaluation period q and (ii) the characteristics of the funding policies including the level of experimentation defined as the ratio of the abandonment budget and the budget for launching projects (c_a/c_ℓ), and the percentage of funding towards large projects.

In Figure 10, we illustrate two of the funding policies presented in Figure 9. The first, more risk-averse funding policy (a) is $\mathcal{P} = (27, 1, 3, 2)$ with a portfolio risk level $\alpha = 0.01$. In the majority of instances, only small projects are launched and 40% of those are launched conditionally. Furthermore, only 30% of these projects are abandoned after the re-evaluation. Thus, the number of completed small projects is large reducing the portfolio risk through diversification. One large project is also launched conditionally in 20% of instances.

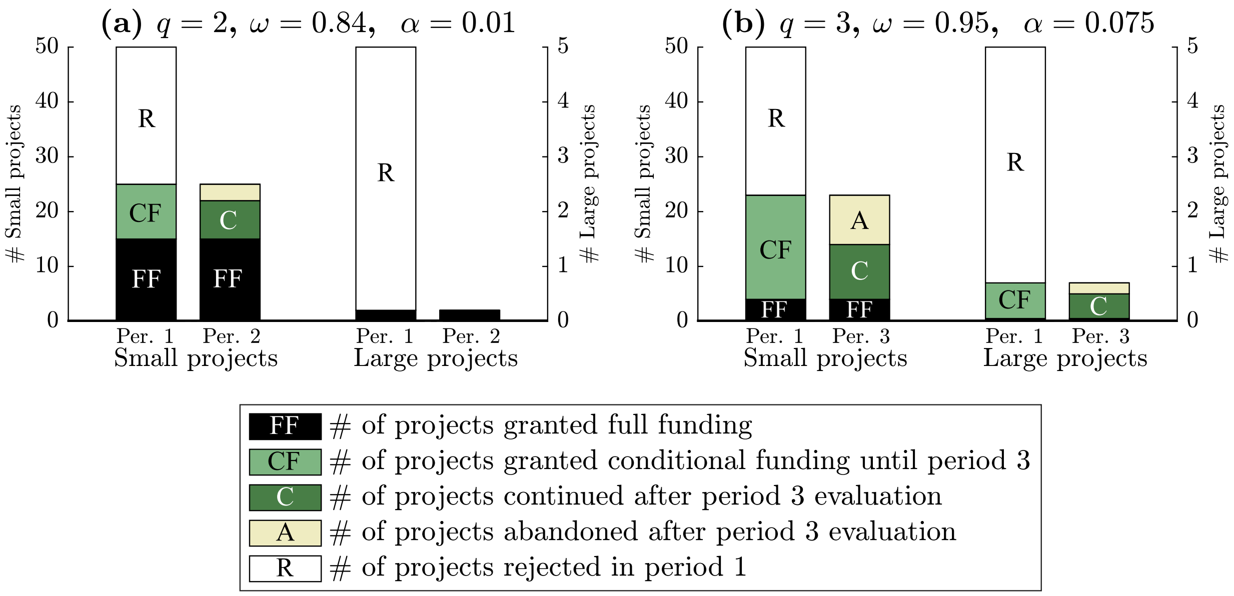

Figure 10: Illustration of two funding policies from the Pareto front with unrestricted re-evaluation period q (Fig. 9) with risk levels (a) $\alpha = 0.01$ and (b) $\alpha = 0.075$.

The second funding policy (b) has a larger portfolio risk level of $\alpha = 0.075$. The funding policy is $\mathcal{P} = (30, 2, 11, 3)$, indicating that more projects are experimented with and the experimentation is carried out longer. Almost all projects are launched conditionally, and, on average, over a third of the launched projects are abandoned after re-evaluation. The funding policy results in 13% more completed excellent projects when compared to funding policy (a). Hence, it seems that more profitable although riskier funding strategies can be obtained by launching more large projects, experimenting with more projects, and conducting the experimentation longer.

5.3 Comparison of alternative project selection strategies

Table 3 describes four alternative funding strategies that we next compare with the Pareto efficient solutions. These strategies specify the stage two and three decisions, i.e., choosing which projects are launched, evaluated, and abandoned. The stage one decision, stable funding policy $(c_\ell, c_\varepsilon, c_a, q)$, is determined for each strategy so that for the portfolio risk level it holds that $\alpha \leq 0.10$, and that the stable funding policy is Pareto efficient given fixed index sets $\mathcal{L}$, $\mathcal{E}$, and $\mathcal{A}$.

Funding strategy	**Avoid large projects**	**Prefer large projects**	**Risk-indifferent one by one selection**	**Risk-informed one by one selection**
Description	Large projects are rejected as too risky	Large projects are preferred	Risk-indifferent funding strategy where projects are chosen one by one	Risk-informed funding strategy where projects' riskiness is considered
Selecting launched projects $\mathcal{L}$	Only small projects are launched	A maximum amount of large projects are launched	Projects are selected one by one to the portfolio according to the resource consumption adjusted excellence probabilities $\mathbb{P}(V_i \geq v_i^\xi)/c_i$	Launched projects are selected similarly as in the risk-indifferent policy except that a 20% risk premium is required for large projects to be selected
Selecting evaluated projects $\mathcal{E}$	Optimal selection	Optimal selection	Similar logic as choosing launched projects	Large projects with $\mathbb{E}[V_i \mid s_i^0] \leq 230$ are evaluated
Selecting abandoned projects $\mathcal{A}$	Optimal selection	Optimal selection	Similar logic as choosing launched projects	Large projects with $\mathbb{E}[V_i \mid s_i^0, s_i^q] \leq 200$ are abandoned

Table 3: Description of alternative funding strategies.

Research agencies typically review and select projects to be funded one by one based on the projects' ranking until the whole funding budget is spent. This ranking can be based on the aggregation of scores of multiple criteria including the riskiness of research projects. However, funding agencies may be too risk-averse when reviewing individual projects as they do not consider the whole portfolio view (Franzoni et al., 2022). The *Avoid large projects* funding strategy attempts to describe aforementioned conditions, where large projects are considered too risky and are always rejected. Thus, only small projects are launched, and they are evaluated and abandoned in an optimal way. An example of this funding strategy was already presented in Figure 4 (a) as the most risk-averse Pareto efficient solution in which all projects were granted full funding. Thus, assessing projects' riskiness and selecting them without considering the whole portfolio view can lead to an undesirably low portfolio risk level.

The *avoid large projects* funding strategy is shown in Figure 11 with a considerably more risk-seeking stable funding policy $(c_\ell, c_\varepsilon, c_a, q) = (32, 2, 15, 3)$. The expected number of completed excellent projects is $\omega = 0.92$, which is only 8% more than with the most risk-averse Pareto efficient solution. The portfolio risk level increases to $\alpha = 0.07$ from near-zero. This stable funding policy allows the experimentation with a larger set of projects, yet those projects are not necessarily of high quality, which is why the benefits of the experimentation are small.

In the *prefer large projects* funding strategy, the decision maker seeks to launch as many large projects as possible, incentivized by the fact that they are more likely to be excellent. In Figure 11 this strategy is shown with stable funding policy $(c_\ell, c_\varepsilon, c_a, q) = (25, 0, 0, 3)$, which means that all projects are granted full funding, and in each period two large and five small projects are launched. The outcome of this strategy is roughly the same as with the *avoid large projects* funding strategy. This shows that both the portfolio level strategy and the selection of individual projects must be aligned to reach optimal solutions.

The third funding strategy, *risk-indifferent one by one selection* attempts to maximize the expected number of completed excellent projects by selecting the projects with highest probabilities of being excellent to the portfolio one by one until the budget is exhausted. This is done with respect to the projects' resource consumption, i.e., a higher quality is expected from project with a higher resource consumption. The same logic is followed when evaluating and abandoning projects except starting from the projects with lowest probabilities. The stable funding policy attached to this strategy is $(c_\ell, c_\varepsilon, c_a, q) = (31, 2, 13, 3)$. This strategy guides the decision maker to launch, on average, 1.4 large and 17 small projects, the majority of which are evaluated. On average, 0.5 large and 8 small projects are abandoned during the re-evaluation period. This funding strategy is quite close to the Pareto front at roughly $\omega = 1.1$ and $\alpha = 0.08$. However, with an optimal funding policy, the same expected number of completed projects could be attainable with roughly half of the risk at $\alpha = 0.04$.

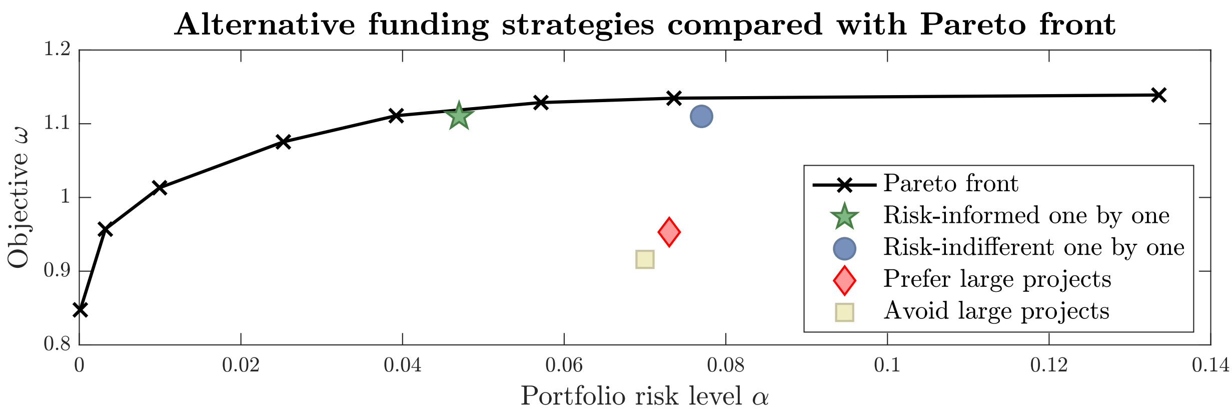

Figure 11: Comparison of alternative funding strategies with Pareto efficient solutions. The funding strategies were simulated using 500×500 Monte Carlo samples.

The *risk-informed one by one selection* strategy attempts to capture the individual projects' risk to reward characteristics by requiring the large projects a 20% larger resource consumption adjusted probability of being excellent when selected to the portfolio instead of a small project. This percentage can be thought as a risk premium required due to the lack of diversification when launching large projects.

In addition, large projects are now evaluated if the initial value estimates $\mathbb{E}[V_i \mid s_i^0]$ are smaller than 230 and are abandoned if the interim value estimates $\mathbb{E}[V_i \mid s_i^0, s_i^q]$ are smaller than 200. This seeks to mimic the phenomena illustrated in Figure 5, where highly promising large projects are committed to from the beginning, majority of funded large projects are granted conditional funding, and evaluated large projects are abandoned quite cautiously. This funding strategy paired up with a stable funding policy $(c_\ell, c_\varepsilon, c_a, q) = (32, 2, 15, 3)$ is near the Pareto front with approximately $\omega = 1.1$ and $\alpha = 0.05$. The main differences to the *risk-indifferent one by one selection* strategy is that large projects are both launched and abandoned slightly less frequently. These observations indicate that all project possibilities should be considered, and both the individual projects' risks and the whole portfolio view need to be taken into account in decision making.

6 Discussion and conclusions

Fostering innovative research has a tremendous positive impact on society through the development of new breakthrough technologies. However, there has been criticism towards public research funding programs due to inefficiency in allocating funds to advance such rare breakthroughs. Some key shortcomings in public research funding are recognised to be, for instance, the lack of a holistic portfolio view in risk management and not funding research projects conditionally in multiple stages. In this thesis, we developed a stochastic optimization framework for forming risk-informed research funding policies to support the development of breakthrough technologies. The framework includes an abandonment option, which grants a possibility to re-evaluate already launched projects and continue only those that seem most promising. Concentration risk in the model is considered by allowing freedom in the projects' sizes and modelled using the Value-at-Risk framework. We described a case problem and solved the Pareto efficient solutions numerically using the sample average approximation method.

Our numerical results support the academic consensus that a holistic portfolio view is important for effective project portfolio selection and management. In our model, the decisions related to project selection can be split into two parts, (i) the selection of the long-term stable funding policy, and (ii) the selection of individual projects in each funding period done in accordance with the funding policy. We found that both decision stages are dependent on the decision maker's risk-appetite, and that considering the interplay of these two stages is crucial for successful decision making. A more risk-seeking funding policy requires a more risk-informed selection of funded projects, and vice versa.

One key consideration in practically all risk management is diversification. In this thesis, we assessed how dominantly large projects affect the performance of a research project portfolio. We found that including large projects to a portfolio of research projects can support the development of breakthrough technologies as a part of an otherwise well diversified portfolio. However, to keep portfolio risk at a desired level, funding large projects should be done cautiously by utilizing conditional funding and a suitable amount of diversification should be ensured by funding smaller projects as well. Conditional funding of large projects reduces the risk of committing a major part of the funding budget to a project which could turn out to perform poorly despite high initial expectations. Only highly promising large project opportunities should be granted full funding without any initial experimentation.

One interpretation of the concept of a large project presented in this thesis is a set of multiple small highly correlated projects. For instance, some fields of research could have individual projects opportunities that are highly correlated due to dependencies on the success of the underlying field or some related technology. Hence, an alternative approach for modelling concentration risk could have been to add different levels of correlation between the projects, inducing concentration risk in a similar fashion as large projects. However, in our model, introducing correlation between the projects would not be straight forward but would require an extensive reformulation of the model framework. Yet another viable approach for modelling risky projects could have been to allow different levels of dispersion in the projects' future value distributions corresponding to different levels of risk. This would have also allowed all the projects to be of the same size, perhaps giving a different point of view on risk-informed research project selection.

In alignment with Vilkkumaa et al. (2015), our numerical results suggest that the option to abandon projects can support the development of breakthrough technologies even as a part of more risk-informed decision maker's funding policy. However, we found evidence that the value of the abandonment option predominantly comes from applying it to a large set of small projects and then continuing only a portion of the most promising ones. On the other hand, the abandonment option seemed to function more as a risk management tool with large projects as most experimented large projects were continued after re-evaluation and only a small number of them were abandoned.

The quality of project opportunities matters. We found that when there is a large amount of promising project opportunities, it is optimal to first experiment with a large portion of those projects and then to continue the best ones. Conversely, when the pool of promising projects is small, it may be better to experiment less and use resources to complete a large number of projects. Our results also suggest that while a highly experimental funding policy may be optimal for developing breakthrough technologies, it may not be optimal from the point of view of promoting science in general as all research efforts accumulate knowledge over time and play an important role as a foundation for future breakthrough technologies. Thus, it is vital for research funding agencies to find suitable balance between promoting innovative research yet also supporting more conventional research endeavours.

We identified two approaches for a decision maker with an increased appetite for risk to promote the development of breakthrough technologies using the abandonment option. First, more project opportunities could be experimented with in general. This increases the likelihood of launching excellent projects that lead to breakthrough technologies. The decision maker could also allocate resources towards experimentation with small projects rather than larger ones, although, this can increase the portfolio risk excessively. Second, the re-evaluation period could be delayed further into the projects' lifecycles to obtain more accurate estimates of the project future values before granting final funding.

Our results suggest that the most risk-averse project funding policy is to complete a maximal number of projects. This is reasonable as the policy provides best diversification. Hence, surprisingly, the option to abandon projects did not reduce portfolio risk. One reason for this is that our model is based on a long-term stable research funding policy with a fixed budget. Thus, the abandonment option obligates the decision maker to abandon a fixed portion of the projects in each period. In more dynamic conditions, the decision maker could, for instance, choose to abandon all launched projects and use the saved budget on the next funding period to launch more promising projects. In such conditions, the abandonment option could be more useful as a pure risk management tool as funds could be steered more efficiently towards higher quality projects. However, this would also require an adaptive budget, i.e., the possibility to save and loan money (see e.g., Tofighian et al. 2018).

We assumed that all projects are successfully completed given that a preknown budget is granted for them. However, this may not hold for real world research projects. For instance, Baker and Solak (2011) provided an optimization framework for selecting a portfolio of R&D projects in response to climate change, in which the success probabilities of the individual projects are dependent on the amount of provided funding. Furthermore, the amount of funding needed to complete a research project may also not be exactly known beforehand, and this should be considered in risk management (see e.g., Hu and Szmerekovsky, 2017). In addition, the projects' future values may depend on aspects such as whether funding is granted for a short time period or for the whole lifetime of the project. Heinze (2008) found that funding programs which grant long-term research funding can encourage risk-taking leading to more breakthrough results and short-term funding can, vice versa, lead to risk-averse research strategies. They also found that big research teams may not be as effective as small ones, indicating that the benefits gained from projects are not always directly proportional to the amount of funding given. Our model does not consider many of such intricacies but is developed on generic assumptions and simplifications. Yet, we believe that we were able to provide valuable insight and principles for selecting a portfolio of risky research projects.

Our numerical results are aligned with the criticism towards research funding programs, showing that there can be a large trade-off between supporting the development of breakthrough technologies and research in a more general sense. Our model shows that conditional project funding can be important in risk-informed development of breakthrough technologies. Yet, conditional funding of small projects did not function as a risk mitigate due to model restrictions. Rather, we found that more risk-seeking funding policies can be formed by increasing the duration and share of the conditional funding. Nevertheless, our results suggest that conditional funding of large projects can be used as a risk management tool while promoting the development of breakthrough technologies.

References

Erin Baker and Senay Solak. Climate change and optimal energy technology R&D policy. *European Journal of Operational Research*, 213(2):442–454, 2011.

Rainer Baule. Allocation of risk capital on an internal market. *European Journal of Operational Research*, 234(1):186–196, 2014.

Philip Best. *Implementing value at risk*. John Wiley & Sons, 2000.

Luigi Buzzacchi. Combining portfolios of risky research projects within a science of science funding. 2022.

Vijaya Dixit and Manoj Kumar Tiwari. Project portfolio selection and scheduling optimization based on risk measure: A conditional value at risk approach. *Annals of Operations Research*, 285(1-2):9–33, 2020.

Ghislaine Elbok and Abdelaziz Berrado. Towards an effective project portfolio selection process. In *International Conference on Industrial Engineering and Operations Management Rabat, Morocco*, 2017.

Chiara Franzoni, Paula Stephan, and Reinhilde Veugelers. Funding risky research. *Entrepreneurship and Innovation Policy and the Economy*, 1(1):103–133, 2022.

Gaia Gasparini, Matteo Brunelli, and Marius Dan Chiriac. Multi-period portfolio decision analysis: A case study in the infrastructure management sector. *Operations Research Perspectives*, 9:100213, 2022.

Banu Gemici-Ozkan, S David Wu, Jeffrey T Linderoth, and Jeffry E Moore. OR PRACTICE — R&D project portfolio analysis for the semiconductor industry. *Operations Research*, 58(6):1548–1563, 2010.

Anna P Goldstein and Michael Kearney. Know when to fold 'em: An empirical description of risk management in public research funding. *Research Policy*, 49(1): 103873, 2020.

Vijay Govindarajan and Praveen K Kopalle. How legacy firms can introduce radical and disruptive innovations: Theoretical and empirical analyses. In *Academy of Management Proceedings*, volume 2004, pages A1–A6. Academy of Management Briarcliff Manor, NY 10510, 2004.

Zvi Griliches. The search for R&D spillovers. *National Bureau of Economic Research Working Paper Series*, (w3768), 1991.

Luca Grilli, Mariana Mazzucato, Michele Meoli, and Giuseppe Scellato. Sowing the seeds of the future: Policies for financing tomorrow's innovations. *Technological Forecasting and Social Change*, 127:1–7, 2018.

Nitish Gupta, Hyunkyu Park, and Rob Phaal. The portfolio planning, implementing, and governing process: An inductive approach. *Technological Forecasting and Social Change*, 180:121652, 2022.

Nicholas G Hall, Daniel Zhuoyu Long, Jin Qi, and Melvyn Sim. Managing underperformance risk in project portfolio selection. *Operations Research*, 63(3):660–675, 2015.

Thomas Heinze. How to sponsor ground-breaking research: A comparison of funding schemes. *Science and Public Policy*, 35(5):302–318, 2008.

Kwon Ryong Hong, Xiaoxia Huang, Jang Su Kim, and Nam Hyok Kim. A multi-objective mean–semivariance model for project selection using reinvestment and synergy under uncertainty. *Expert Systems with Applications*, 217:119586, 2023.

Qiaohai Hu and Joseph Szmerekovsky. Project portfolio selection: A newsvendor approach. *Decision Sciences*, 48(1):176–199, 2017.

Xiaoxia Huang, Lan Xiang, and Sardar MN Islam. Optimal project adjustment and selection. *Economic Modelling*, 36:391–397, 2014.

Matthijs J Janssen. What bangs for your buck? Assessing the design and impact of dutch transformative policy. *Technological Forecasting and Social Change*, 138: 78–94, 2019.

Janne Kettunen and Ahti Salo. Estimation of downside risks in project portfolio selection. *Production and Operations Management*, 26(10):1839–1853, 2017.

Anton J Kleywegt, Alexander Shapiro, and Tito Homem-de Mello. The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, 12(2):479–502, 2002.

Ronald Klingebiel and Christian Rammer. Resource allocation flexibility for innovation performance: The effects of breadth, uncertainty, and selectiveness. *ZEW Discussion Papers*, No. 11-073, 2011.

Josh Lerner. *Boulevard of broken dreams: Why public efforts to boost entrepreneurship and venture capital have failed - and what to do about it.* Princeton University Press, 2009.

Josh Lerner and Ramana Nanda. Venture capital's role in financing innovation: What we know and how much we still need to learn. *Journal of Economic Perspectives*, 34(3):237–261, 2020.

Matthew Li, Jun Lu, Zhongwei Chen, and Khalil Amine. 30 years of lithium-ion batteries. *Advanced Materials*, 30(33):1800561, 2018.

Juuso Liesiö, Ahti Salo, Jeffrey M Keisler, and Alec Morton. Portfolio decision analysis: Recent developments and future prospects. *European Journal of Operational Research*, 293(3):811–825, 2021.

Gustavo Manso. Motivating innovation. *The Journal of Finance*, 66(5):1823–1860, 2011.

Guilherme Augusto Barucke Marcondes. Project portfolio selection considering return-risk evaluation and multiple-criteria decision analysis. In *ICORES*, pages 264–269, 2019.

Harry Markowitz. Modern portfolio theory. *The Journal of Finance*, 7(11):77–91, 1952.

Camilo Micán, Gabriela Fernandes, and Madalena Araújo. Project portfolio risk management: A structured literature review with future directions for research. *International Journal of Information Systems and Project Management*, 8(3):67–84, 2020.

CM Mottley and RD Newton. The selection of projects for industrial research. *Operations Research*, 7(6):740–751, 1959.

Gina Colarelli O'Connor, Thiagarajan Ravichandran, and Daniel Robeson. Risk management through learning: Management practices for radical innovation success. *The Journal of High Technology Management Research*, 19(1):70–82, 2008.

Gunnar Öquist and Mats Benner. *Fostering breakthrough research: A comparative study*. Kungl. vetenskapsakademien Stockholm, 2012.

Research Council of Finland. Biofuture2025 (2017-2025). https://www.aka.fi/en/research-funding/programmes-and-other-funding-schemes/academy-programmes/biofuture2025-2017-2021/, 2017. Accessed: 2024-02-25.

Kevin Roberts and Martin L Weitzman. Funding criteria for research, development, and exploration projects. *Econometrica: Journal of the Econometric Society*, pages 1261–1288, 1981.

R Tyrrell Rockafellar and Roger J-B Wets. *Variational analysis*, volume 317, chapter 7, pages 238–269. Springer Science & Business Media, 2009.

Anna Le Gerstrøm Rode, Per Svejvig, and Miia Martinsuo. Developing a multidimensional conception of project evaluation to improve projects. *Project Management Journal*, 53(4):416–432, 2022.

Sheldon M Ross. *A first course in probability*, pages 360, 400. Pearson Boston, 8. edition, 2014.

Ahti Salo, Jeffrey Keisler, and Alec Morton. *Portfolio decision analysis: Improved methods for resource allocation*, volume 162. Springer, 2011.

Ahti Salo, Michalis Doumpos, Juuso Liesiö, and Constantin Zopounidis. Fifty years of portfolio optimization. *European Journal of Operational Research*, 2023. ISSN 0377-2217. doi: https://doi.org/10.1016/j.ejor.2023.12.031.

Leonardo P Santiago and Pirooz Vakili. On the value of flexibility in R&D projects. *Management Science*, 51(8):1206–1218, 2005.

Samantha Sharpe, Andy Cosh, and David Connell. Funding breakthrough technology. *Cambridge Integrated Knowledge Centre Report, Cambridge*, 2013.

James E Smith and Robert L Winkler. The optimizer's curse: Skepticism and postdecision surprise in decision analysis. *Management Science*, 52(3):311–322, 2006.

Gerard J Tellis, Jaideep C Prabhu, and Rajesh K Chandy. Radical innovation across nations: The preeminence of corporate culture. *Journal of Marketing*, 73(1):3–23, 2009.

Ali Asghar Tofighian, Hamid Moezzi, Morteza Khakzar Barfuei, and Mahmood Shafiee. Multi-period project portfolio selection under risk considerations and stochastic income. *Journal of Industrial Engineering International*, 14:571–584, 2018.

Eeva Vilkkumaa, Ahti Salo, Juuso Liesiö, and Afzal Siddiqui. Fostering breakthrough technologies — How do optimal funding decisions depend on evaluation accuracy? *Technological Forecasting and Social Change*, 96:173–190, 2015.

Steven R Walk. Quantitative technology forecasting techniques. *Technological Change*, 24, 2012.

Shlomo Yitzhaki and Edna Schechtman. *The Gini methodology: A primer on a statistical methodology*, volume 272. Springer, 2013.