



CZECH TECHNICAL UNIVERSITY IN PRAGUE
Faculty of Nuclear Sciences and Physical Engineering



Real Options Valuation: A Dynamic Programming Approach

Oceňování projektů metodou reálných opcí z pohledu dynamického programování

Master's Thesis

Author: **Filip Rolenec**
Supervisor: **Ing. Rudolf Kulhavý, DrSc.**
Language advisor: **Ing. Rudolf Kulhavý, DrSc.**
Academic year: 2020/2021

ZADÁNÍ DIPLOMOVÉ PRÁCE

Student:	Bc. Filip Roleneč
Studijní program:	Aplikace přírodních věd
Obor:	Matematické inženýrství
Název práce (česky):	Oceňování projektů metodou reálných opcí z pohledu dynamického programování
Název práce (anglicky):	Real Options Valuation: A Dynamic Programming Approach

Pokyny pro vypracování:

1. Seznamte se s tradičním přístupem k analýze reálných opcí obvyklým ve finanční analýze.
2. Formulujte analýzu reálných opcí jako úlohu stochastického řízení.
3. Navrhněte vhodnou metodu numerické aproximace dynamického programování.
4. Implementujte algoritmus oceňování ve Vámi zvoleném výpočetním nástroji a demonstруйте jeho chování na ilustrativní aplikaci a simulovaných datech.
5. Analyzujte přínosy teorie stochastického řízení pro analýzu reálných opcí. Identifikujte případná omezení a otevřené otázky.

Doporučená literatura:

1. Copeland, Thomas E., and Vladimir Antikarov, Real Options: A Practitioner's Guide. Revised ed. New York: Texere, 2003.
2. Guthrie, Graeme, Real Options in Theory and Practice. Oxford, England: Oxford University Press, 2009.
3. Powell, Warren B, Approximate Dynamic Programming: Solving the Curses of Dimensionality. 2nd ed. Hoboken, NJ: Wiley, 2011.
4. Puterman, Martin L., Markov Decision Processes: Discrete Stochastic Dynamic Programming. Hoboken, NJ: Wiley, 2005.
5. Vollert, Alexander. A Stochastic Control Framework for Real Options in Strategic Valuation. Boston, MA: Birkhäuser, 2003.

Jméno a pracoviště vedoucí diplomové práce:

Ing. Rudolf Kulhavý, DrSc.

Batličkova 253/1, , 182 00 Praha 8

Jméno a pracoviště konzultanta:

Datum zadání diplomové práce: 31.10.2019

Datum odevzdání diplomové práce: 4.5.2020

Doba platnosti zadání je dva roky od data zadání.

V Praze dne 11. října 2019

.....
garant oboru

.....
vedoucí katedry

.....
děkan

Acknowledgment:

I would like to thank my supervisor Ing. Rudolf Kulhavý, DrSc. for his professional guidance and all the advice given while creating this thesis.

Author's declaration:

I declare that this Master's thesis is entirely my own work and I have listed all the used sources in the bibliography.

Prague, May 7, 2020

Filip Rolenec

Název práce:

Oceňování projektů metodou reálných opcí z pohledu dynamického programování

Autor: Filip Rolenic

Obor: Matematické inženýrství

Druh práce: Diplomová práce

Vedoucí práce: Ing. Rudolf Kulhavý, DrSc.

Abstrakt: Oceňování převážné většiny investičních příležitostí se dnes stále určuje metodou *diskontovaných peněžních toků* (DCF) ???. DCF metoda je přímočará a pro jednoduché projekty dává velmi přesné výsledky. Složitější projekty, které pro tuto práci definujeme jako projekty s vysokou mírou vnitřní neurčitosti a existencí manažerských rozhodnutí značně ovlivňujících strukturu projektu (reálné opce), jsou dnes oceňovány metodou *real option analysis* (ROA). Metoda ROA přiznává hodnotu možnostem změny projektového plánu, v důsledku čehož je ohodnocení ROA vyšší než DCF.

Tato práce má za cíl interpretovat řízení projektů a s ním související ocenění, jako řízení stochastického systému. Rozsáhlá teorie stochastického řízení ??, ??, ??, dovoluje vybudovat novou teorii oceňování postavenou na základech ROA, která mimo jiné řeší omezení ROA na pouze jediný zdroj neurčitosti - například cenu těžebních minerálů na trhu [].

Nový originální přístup k oceňování projektů, který kombinuje dosavadní dosažené znalosti v ekonomické teorii (ROA) a teorii stochastického řízení systémů (SDT) je obecně definován v první části práce a prakticky otestován simulacemi na třídě problémů z oblasti <Area of study>.

Simulace ukazují, že kromě lepší interpretovatelnosti modelu a možnosti vložení subjektivních pravděpodobností je ocenění metodou ... přesnější o ... %. Výsledky vyjadřují naději na zvýšenou adopci metody ... v praxi.

Klíčová slova: Analýza reálných opcí, Diskontované peněžní toky, Oceňování projektů, Stochastické řízení, Těžební průmysl

Title:

Real Options Valuation: A Dynamic Programming Approach

Author: Filip Rolenic

Abstract: The valuation of investment opportunities (projects) is nowadays still predominantly computed by the *discounted cash flow* (DCF) method ???. DCF is straightforward and gives solid results for simple projects. For more complicated projects, which are in this thesis defined as projects with substantial degree of inner uncertainty and with an existence of ability to change the course of the project by direct managerial actions, the *real option analysis* (ROA) is now being used. ROA recognizes the value in the possibility of projects' plan change, which usually results in loss limitation, increasing the project's overall expected value.

This thesis aims to interpret project execution and its valuation as a problem of stochastic decision control. The extensive stochastic decision control theory (SDT) allows, based on ROA, to design new valuation theory, that for example solves the limitation of ROA on only one source of uncertainty - such as the volatile price of the mined materials on the market.

The new and original approach to project valuation, that combines both the economical theory (ROA) and the stochastic control theory is defined in the first part of the thesis and then tested by simulations on a class of valuation problems - valuation of projects in <Industry>.

Simulations show that in addition to better model's interpretability and a possibility to include a personal probability distributions, the valuation by <method> is <percentage> more precise than best current strategies. The results hint a possible increase in adoption of this method in practice.

Keywords: Discounted cash flow, Mining industry, Project valuation, Real option analysis, Stochastic decision control

Contents

1	Introduction	17
2	Valuation task	21
2.1	How is a similar task solved by different authors	22
2.2	My valuation	23
2.2.1	Addressing the uncertainty	23
2.2.2	Actual value computation	24
2.2.3	The maximization process	27
2.2.4	Note in the end	27
3	Preliminaries	29
3.1	Classic valuation techniques	29
3.1.1	Net Present Value	29
3.1.2	ENPV valuation	30
3.1.3	Comparing investments	30
3.1.4	Other frequently used valuation techniques	30
3.1.5	Decision tree analysis	30
3.2	Real option analysis	30
3.2.1	Financial options	31
3.2.2	Black-Scholes Merton model	31
3.2.3	Real options	31
3.3	Statistical decision theory	31
3.3.1	Bayesian statistics	32
3.3.2	Utility	32
3.4	Approximate dynamic programming	33
4	Real option analysis in terminology of stochastic decision control	35
5	Application of the cheery-picked valuation technique.	37
5.1	Approximate dynamic programming	37
5.2	Valuation example	37
5.2.1	NPV valuation	37
5.2.2	DTA valuation	37
5.2.3	Classic ROA valuation	37
5.2.4	New approach	37
6	Discussion	39

7 Conclusions 41

8 Bachelor’s thesis parts that could be useful for TeX styling 43

8.1 Discrete Markov decisions processes 43

Chapter 1

Introduction

The current state of capital investment valuation techniques is a mess. Majority of investors still rely on the standard Net Present Value (NPV) technique or its slight generalizations in form of optimal/average/bad scenarios, risk adjusted discount rate (RADR) hurdles (usually set to an arbitrary percentage), internal rate of return (IRR) metric (which is essentially the same as RADR hurdles) and other metrics like,... ROIC, ENPV.

All the listed generalizations try to cope with one of two main problems of NPV, which is inability to incorporate uncertainty (ENPV, scenario approach) or intra-investment comparison (IRR(?), ROIC).

None of these valuation techniques acknowledges the managerial ability to take action and improve the course of an investment. ENPV can be interpreted is such way but that is abusing of its notation.

The decision tree analysis (DTA) is the simplest approach for valuing an investment with acknowledging decision nodes as an inherit trait of an investment process. It should not be a surprise that being able to make a decision through the lifetime of an investment has value. Also, it is quite clear that the more uncertain the investment and its parts are, the more valuable the ability to act is. To value an ability to make a decision and change the path which the investment follows is not a complicated task. All it takes is to compare a value of a project with that ability and without it (both with DTA), and declare that the value of such "decision option" is the difference between the two.

Many authors from the financial world ¹ use the term real option valuation when talking about this difference between DTA trees with or without some branches that represent the ability to make a decision and act upon it. I think that is rather unfortunate, since the terminology clashes with the real option valuation, which this thesis is about.

The true real option analysis (ROA) is about being able to value investment options in the same way financial options are valued. The Black-Scholes-Merton option pricing model that is widely used in financial world is an exceptionally elegant valuation tool. All that it needs to know about financial option to value it is four parameters - the strike price and expiration date of an option itself, the volatility and current price of the underlying asset upon which the option is written. ². This Nobel-Prize winning elegance led mathematical and financial experts to try to interpret investment decision making in terms of "real options", that can be valued in the same way the financial options are, by the BSM model.

The attempts to interpret each of four important parameters of a financial option in terms of investment parameters are successful for simple cases (Dealership in BERK), however they fail for more complex ones. The ability to use real option valuation technique depends heavily on the ability to find a relevant company, stock or other publicly traded proxy variable to the examined investment. One finds plausible to construct a replicating portfolio for an investment in oil mining facility or a food production

¹ See *this* and *that*

² probably some more assumptions

line, however this assumption is hard to believe for highly innovative products in new fields, such as software development or R&D investments.

The ROA valuation approach in investment decision making has cyclical nature. The first wave of the real ROA approach can be traced to ... who presented his ideas shortly after the initial publication about financial option pricing [3]. Next wave comes with ... and finally the most recent publications trying to make practical use of real option analysis are [] and [].

Nearly all publications of real options published in the last <number> years talk about small penetrability of this approach to the actual usage by managers in large and small investment companies. The stated reasons are always the same,

- Vollert says that its complexity, but that is because he uses stochastic models, Ito lemma and computes PDEs, all the time. That has nothing to do with replicating portfolios and BSM model.
- Mr. Kulhavy says that the problem is in finding the replicating portfolio and even if you find it you have to persuade the manager to see it too.
- <Problem with adoption>
- <Another problem with adoption>

I believe that the current usage of ROA by investors makes sense. The actual ROA approach is based on finding a replicating portfolio that behaves in the same way as the investment. If a manager believes that he has found a replicating portfolio, then it does not make sense for him to value the project differently than the replicating portfolio. However, there is still the issue of finding such replicating portfolio.

The usage of ROA in sense of different DTA trees is well posed. The ability to hold production in refinery in times of large oil prices can surely increase its value. In this sense the ROA has the biggest impact on the valuation in investments that are highly uncertain and in which there are actions that can alter the course of future cash flow. This also means that in the opposite case, when there is high degree of certainty about the execution and cash flow of the project, or if there are no actions to be taken, the ROA does not create additional value.

The power of classic ROA is to seemingly eliminate the probabilistic nature from the picture. This can be done by the power of all-wise market, that forbids the existence of arbitrage. If a tradeable³ replicating portfolio is found, then I should not care if I make my investment or if I use the money to buy the replicating portfolio.

For the mathematicians from the field of stochastic decision making the problem of valuation is just another decision making under uncertainty. Define state and action space, define rewards (free cash flow) in each epoch and get prior probability density as expected transition probability from one state to another upon undertaking certain action.

The difference between the SDT and ROA is that the former relies on the prior density functions, that are the core of Bayesian statistic, while the latter, coming from economics, relies on the power of the market and managers ability to find a replicating portfolio.

In this thesis the state-of-the art theories of ROA valuation and SDT are presented. Their strengths and weaknesses are discussed and a hybrid approach from the cherry-picked parts of the corresponding theories is presented.

The power of the new theory is shown on a class of problems called ... The current literature copes with them in much simpler way and their assumptions are much larger. My⁴ approach also potentiates adoption in practical usage, as it puts the decision making manager in charge. The manager decides

³This is a valid word, Cambridge dictionary say so...

⁴Or the presented approach, should I avoid pronouns in the text?

how to create the prior distributions, he is the one who needs to determine all the possible actions and scenarios that can happen. The algorithm that is presented in this thesis can and should be used in real world as a managerial tool to maximize the profits of divisions or companies as a whole.

The goal to maximize wealth in financial world is approached in a really narrow matter. Money represents value, which in models decreases in time by a percentage given by some authority, most of the time central banks. In respected financial publications [2], [] or [], there is no notion of utility. It seems like for financial institutions, it does not matter to get 1M in cash now or to gamble for 2M in a fair coin flip. On the other hand, in SDT, the concept of utility is taken very seriously. Majority of people would prefer even a very small amount, such as 200k against a fair gamble between 0 and 2M. This is due to the concave utility function people tend to have when very high amounts are discussed and the question is more about what would make them happier.

Maximization of wealth either in the sense of investment company or individual is natural. Almost everybody would agree that to have more wealth is better than have less. Better investment decisions mean higher accumulation of wealth, which together with assumption of free market results in an increase of utility of each participant in the economy ⁵. In other words, when a company makes higher profits it should mean that the part of the population that participated in the wealth creation should be happier.

⁵Find that reference... Duke university lectures...

Chapter 2

Valuation task

The task is to value a project in a form of $\max E\{NPV(\text{Options})\}$.

Definition 2.1. *A project is defined as a piece of planned work or an activity that is finished over a period of time and intended to achieve a particular purpose, mainly an increase of company's or individual's wealth¹.*

Examples of a project are:

- developing a copper mine;
- innovation of chemical processes in an oil refinery;
- upgrade of current machinery in a production line;
- changing the form of software development philosophy towards agile practices.

When examining the expression $\max E\{NPV(\text{Options})\}$ four observations come to mind.

First, the maximization is over some set of control strategies. This set can be generally large, even uncountable. Due to the class of the problems that are addressed in this thesis, the actions made by a manager are not expected to be continuous, not even very frequent. This would result in an assumption of small control strategy space, at least for now.

Assumption 1. *The strategy space is small, certainly finite.*

Second, there is an expectation E , which means that there is uncertainty for a single fixed strategy. This uncertainty comes from the nature of free cash flow in each period and originates in both project's inner development (Will I be able to change the production line to produce the goods I want it to?) or outer, most of the time in a form of market price uncertainty of input or output goods.

Definition 2.2. *Free cash flow is the cash a company produces through its operations, less the cost of expenditures on assets.*

Third, net present value (NPV), that is defined in all publications that I have seen as:

$$NPV = \sum_{t \in T} \frac{C_t}{r^t}, \quad (2.1)$$

where C_t is free cash flow and r is a discount rate that represents devaluation of money in time.

The last observation is that NPV is a function of options. There has to exist some opportunity to change future cash flow in the light of partial or full uncertainty resolution.

¹First part comes from Cambridge dictionary.

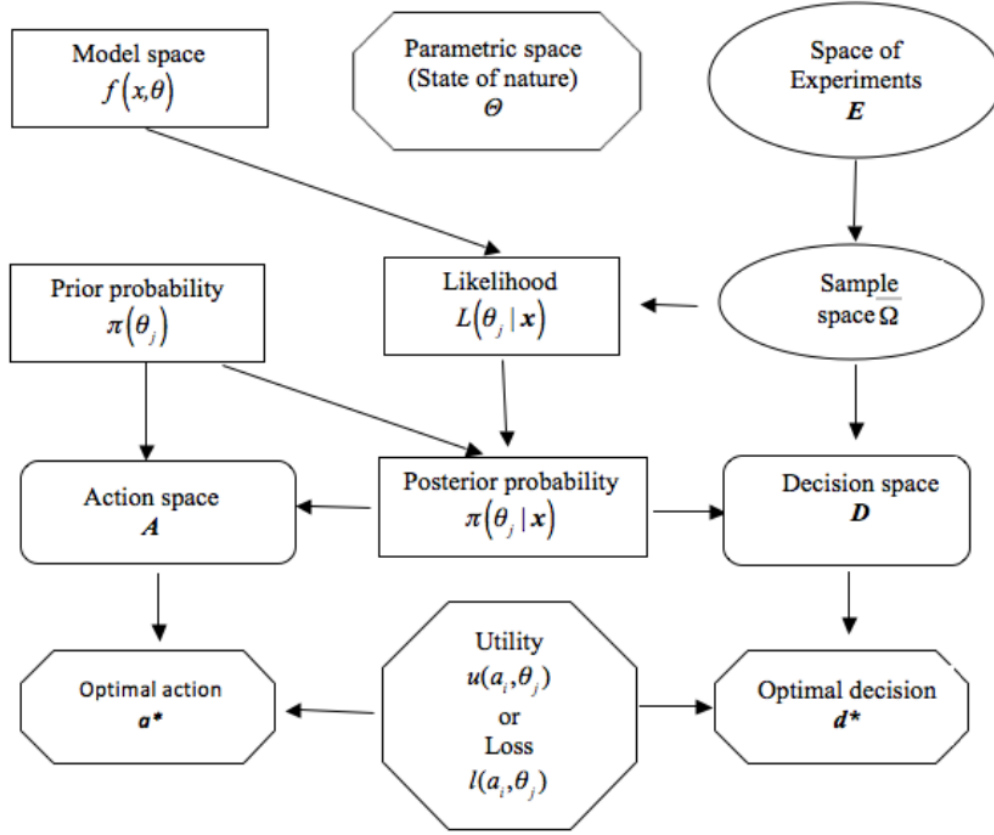


FIGURE 5.1
Conceptual frame of a decision-making process.

Fig. 2.1. SDT Schema.

2.1 How is a similar task solved by different authors

Statistical decision theory is in my opinion captured the best in a book from Bacci and Chiandotto [1] and summarized by Fig. 2.1.

The project (SDT authors call it a *process*) is a sequence of states and actions (and sometimes decisions). To obtain the best performance out of the project a measure of utility instead of simple monetary reward can be introduced. The uncertainty in a project is addressed by a parameterized model with prior distribution that concentrates all knowledge relevant to the project.

The best action strategy is the one that maximizes cumulative utility and it is usually obtained by a method of backward induction called dynamic programming.

Definition 2.3. *TODO - a proper definition of dynamic programming.*

Not sure about what is the difference between stochastic decision theory and stochastic control theory. They are probably synonyms.

Dynamic programming is a tool used for determining the optimal strategy and a value of a project (*process*). Its core is the Bellman equation, which represents the backward induction of value function computation.

Real option analysis can be understood in multiple ways. If a replicating portfolio exists, then a value of some expected cash flow is identical to the current value of a portfolio. The support for this statement comes from the law of one price from economics. The assumption of replicating portfolio existence becomes hard to defend as the project that is to be valued becomes more abstract and/or complex.

For more complex projects, real option valuation uses binomial models with risk-neutral probabilities coming from the law of one price (Guthrie [5]), or stochastic models of an expected behavior of the state variable determining the future cash flow (Vollert [6]).

2.2 My valuation

This sub-chapter is my attempt to best describe a real option valuation understood in a sense of an opportunity to make change in future project's cash flows.

First I start with the intuitive equation:

$$V = \max_{\pi \in \Pi} E \left[\sum_{t \in \mathbf{T}} DC_t(s_t, b_t(\pi)) \right], \quad (2.2)$$

where π is a strategy, a sequence of decisions that the manager is able to do to change the course of the project. Time evolution of the process is represented by a variable t which iterates over range of time epochs, so far finite or infinite countable \mathbf{T} . Discounted cash flow DC_t assumed to be obtained at the end of an epoch t is derived from the free cash flow C_t and a constant discount rate $r > 1$ as:

$$DC_t = \frac{C_t}{r^t}. \quad (2.3)$$

Assumption 2. *Discount parameter is constant in time.*

We can see that DC_t depends on two variables, state variable s_t , which represents the state of an environment (outer world, like prices of oil, price of skilled software engineers, etc.) and b_t , project "branch", which represents that state of the project itself (inner property of the project, for example if a copper mine is active or idle, or if the decision was to produce chocolate instead of vanilla ice cream).

Assumption 3. *State variable s_t is a random variable and does not depend on actions of the project's management.*

Assumption 3 can be violated for example in duopolistic markets or others, where a single company decision has non-negligible influence on the market.

Assumption 4. *A "branch" upon which the project is in the moment is assumed to be a random variable and its value is determined only by actions in a policy π and its random nature.*

Assumption 4 is reasonable for majority of projects. Management makes a decision to open or close a mine, to produce chocolate or vanilla. However, for example in some R&D projects the outcome of a manager's decision is uncertain.

2.2.1 Addressing the uncertainty

In this subsection the meaning of the expected value will be examined. We follow on the equation 2.2². Using the linear property of expectation E and substituting for DC_t with equation 2.3 we obtain:

²not sure about the expression

$$V = \max_{\pi \in \Pi} \sum_{t \in \mathbf{T}} \frac{E[C_t(s_t, b_t(\pi))]}{r^t}. \quad (2.4)$$

Assumption 5. Environment s_t can take on values from a finite or countable set \mathbf{S}_t , $\forall t \in \mathbf{T}$.

This assumption is important only for simpler discussion later. Uncountable \mathbf{S}_t could probably work too, but we would need to perform integration instead of summation³.

Assumption 6. The actual state of $b_t \in \mathbf{B}_t$ depends only on actions in policy π that happened before the time period $t \in \mathbf{T}$.

As a result of assumption 6 a term *partial strategy* is defined.

Definition 2.4. A partial strategy π_t for a full strategy π is defined as the first t elements of its vector, $\pi_t = (a_1, a_2, \dots, a_t)$ ⁴.

Assumption 7. For a given partial strategy π_t , the project "branch" can take on values from a finite set \mathbf{B}_t , $\forall t \in \mathbf{T}$.

In most cases the "branch" of a project is a deterministic function of a managerial decision, however b_t also covers variables whose evolution is triggered by a managerial decision. For example building a bridge with toll payments a year later also postpones the "adoption curve" of the new road.

Assumption 8. The state of an environment s_t and the branch state b_t are stochastically independent random variables.

Using assumptions 6, 8 and definition 2.4 we arrive at the final valuation formula:

$$V = \max_{\pi_{t-1} \in \Pi} \sum_{t \in \mathbf{T}} \frac{\sum_{s_t \in \mathbf{S}_t} \sum_{b_t \in \mathbf{B}_t(\pi_{t-1})} p(s_t) p(b_t(\pi_{t-1})) C_t(s_t, b_t(\pi_{t-1}))}{r^t}. \quad (2.5)$$

Now, there are several questions about our ability to actually compute this valuation.

- Where do the sets \mathbf{T} , \mathbf{S}_t , \mathbf{B}_t come from?
- How to determine the discount factor r .
- How are we able to obtain the cash flow function $C_t(s_t, b_t)$?
- What are we maximizing over? What is the set of possible strategies $\pi \in \Pi$?
- How to obtain the probabilities $p(s_t)$, $p(b_t(\pi_{t-1}))$?

All of these questions will be answered in the next section.

2.2.2 Actual value computation

In this section we will be discussing, how to actually value a project with the equation 2.5.

³The verb integrate probably has a different meaning.

⁴It can be noticed that $\pi_{|\mathbf{T}=\pi}$

2.2.2.1 Sets \mathbf{T} , \mathbf{S}_t , \mathbf{B}_t

All of the sets \mathbf{T} , \mathbf{S}_t , \mathbf{B}_t come from the model of a valuation problem.

The number and frequency of time periods is determined by \mathbf{T} , when the most usual approach is to have \mathbf{T} as a finite number of yearly periods. This can represent either a duration of project rights (for example in mining) or rights for a cash flow share (for example from bridge tolls), or such fuzzy future, where nobody really knows what will happen with the project 50 years from now.

The set of possible branches \mathbf{B}_t comes from the project manager. He needs to identify the project alteration possibilities in each time period $t \in \mathbf{T}$. The set \mathbf{B}_t is small most of the time, however in cases, when for example a start of a variable that is modeled by binomial tree is dependent on the strategy, the set \mathbf{B}_t can grow rather quickly. The example is again a bridge project, where the number of toll-paying customers in time t depends on a time period in which the bridge was finished.

And finally, the values of an environment state \mathbf{S}_t are determined by a model of the outer world. The set \mathbf{S}_t should be able to capture the uncertainty about all important parameters for the project while remaining as simple as possible. A classical example is a binomial tree of input prices for our process (oil, software engineer's salary).

2.2.2.2 Discount factor r

The discount factor, representing time value of money plays a big role in the financial world. It is expected to be $r > 1$ because today's economy depends on non-negative inflation.

Assumption 9. *The discount factor r is expected to be larger than one and constant through the project's lifetime.*

The assumption 9 serves as a simplification. To study time variant discount factor in the light of real option valuation is not interesting. However, one could simply model the discount rate as r_t with some probabilities $p(r_t)$ and compute an expected discount rate.

2.2.2.3 Cash flow function

The core of each project valuation is its cash flow function $C_t(s_t, b_t)$, which represents the additional monetary value created by it over some time period. Usually, this function is simply defined as the difference between the total price of inputs (usually given by s_t) and outputs (usually given by b_t or s_t) for a given period. In more complex models, the inputs can also represent depreciation of machinery, expenses on marketing campaigns or additional costs to acquiring new workforce (not only their salary).

An example of zero cash flow is when a project is deferred and waits to be undertaken or not. A simple example of non-zero cash flow is a project of running ammonia factory, where the cash flow is:

$$C_t = \frac{n_g}{c_{g \rightarrow a}} p_a - n_g * p_g, \quad (2.6)$$

where n_g is the volume of processed gas, p_g, p_a are prices of gas and ammonia on the market and $c_{g \rightarrow a}$ is an inner company coefficient, that says how many units of gas are needed to produce one unit of ammonia.

2.2.2.4 Strategies π

As was pointed out by many articles ⁵ the ability to change the course of a project has value. The sequence of actions (in some papers called decisions) form a strategy $\pi = (a_1, a_2, \dots, a_{|\mathbf{T}|})$.

⁵See for example [], [], or [].

Assumption 10. Action a_t is expected to happen in the end of the corresponding time period $t \in \mathbf{T}$.

The set of possible strategies Π needs to be given by a manager responsible for the project or his team. He should know best, what are the possible changes in the project, that could improve its cash flow and increase its value.

The set of possible strategies is usually well defined by a real options understood in a sense of an ability to act.

In an example from [5], an operator of a gas power plant can choose to produce electricity or not to, based on the current price of gas and power. The strategies in this example can be represented as binary vectors:

$$\Pi = \{(a_1, \dots, a_{|\mathbf{T}|}) | a_i \in \{1, 0\}, \forall i \in \mathbf{T}\}, \quad (2.7)$$

where $a_i = 1$ means that the plant is running in the next time period $i + 1$ and $a_i = 0$ means that the power plant will be in an idle state.

2.2.2.5 Probabilities

Finally, the most complex problem from the list in the end of section 2.2.1 is to determine what to substitute all $p(s_t)$ and $p(b_t)$ in the equation 2.5 for.

The difference in the ROA, SDT and for example DTA techniques is mainly in the difference in which they cope with the derivation of these values.

For DTA, both $p(s_t)$ and $p(b_t)$ are assumed to be known in each time period. This is because problems solved with DTA are usually simple and they need to declare only small number of probabilities.

For ROA the probabilities $p(s_t)$ arise as the so-called risk-free probabilities that come from the idea of replicating portfolio and the validity of the law of one price. For example Guthrie in [5] uses the capital asset pricing model (CAPM), that provides risk-neutral probabilities of up and down movements, p_u, p_d . All other probabilities that could be defined as $p(s_t)$ are in [5] determined by a binomial pricing model as a product of p_u and p_d .

Another approach to ROA can be seen in the work of Vollert [6], who models the replicating portfolios as some stochastic Ito process, which gives the probabilities of s_t .

The branch probabilities $p(b_t)$ are in ROA not understood as a separate category, thus they do not need to be further discussed⁶.

Finally, in SDT, the value of both $p(s_t)$ and $p(b_t)$ comes from the knowledge about the underlying problem expressed as prior (or posterior, when updated for new data) probability distribution. The prior probability distribution is constructed from the inner knowledge of experts on the subject, former relevant data, principle of insufficient reasons or their combinations. In SDT it is usually expected that $p(s_t)$ comes from some parameterized family of distributions:

$$\mathcal{F} = \{p(s_t|\theta) | \theta \in \Theta\}. \quad (2.8)$$

An example is a binomial distribution $Bi(t, \theta)$:

$$p(s_t = k|\theta) = \binom{t}{k} \theta^k (1 - \theta)^{t-k}. \quad (2.9)$$

I believe that this approach is the most general, because one can interpret the prior distribution in many possible ways. The SDT approach also allows for combination of many different sources of knowledge about the project [].

⁶Not sure about this, check it.

The problems of ROA approach are the inability to find a replicating portfolio to any more complex project and unclear and undiscussed origins of "actual" probabilities in CAPM model and other variables in it used in [5].

The STD approach is protected against these problems, as even when there is no prior, there is always a principle of insufficient reasons, which needs only a model, a requirement of all techniques that try to determine $p(s_t)$ and $p(b_t)$.

2.2.3 The maximization process

TODO after supervisor's review. In a nutshell it is a dynamic programming and after realization that the complexity is large, it is approximate dynamic programming.

2.2.4 Note at the end

Without using the replicating portfolio, the "Real Options" means only the ability to change the projects path as is presented in articles [4],... I would love to incorporate the law of one price in the priors of SDT somehow.

Chapter 3

Preliminaries

To understand the main message of this thesis it is important to present several key financial and mathematical concepts. Two main concepts that will be examined in this thesis are Real Option analysis (ROA) and stochastic decision theory (SDT).

3.1 Classic valuation techniques

For clarity of the further research let's define the basic valuation techniques. Even though, or maybe exactly because, they are simple, they are widely used in investment companies today. All techniques are based on an idea of cash flow per sample period, where cash flow is clear profit, already adjusted for expenditures, taxes, and all other mathematically distracting aspects of accounting.

3.1.1 Net Present Value

This metric is mentioned in almost all publications about valuation techniques. It is simple and it takes into account time value of money. The classic definition of NPV is:

$$NPV = \sum_{t \in \mathbf{T}} \frac{C_t}{r^t}, \quad (3.1)$$

where \mathbf{T} is set of time periods in which free cash flow C_t is obtained. Discounting factor r^t expresses the time value of money and it is expected to be positive. In the light of recent events in advanced economies like that of Japan, Switzerland and Germany, the positivity of this variable is not a clear cut. Unprecedented moves of the central banks of these countries to boost the economy by negative interest rates might cause a problems in NPV valuation, since many examples [],... rely on the fact that an infinite sum of constant cash flows discounted by any factor $r > 1$ converges.

However that is not the concern of this publication and we will assume that $r > 1$.

The NPV valuation technique is simple to use, assuming we know the discount rate, which is constant through the lifetime of the investment, and free cash flows, that are expected to be certain.

The problem with this valuation arises if we acknowledge that the discount rate, or cash flows, or both can be uncertain. Two possible valuation techniques emerge from this uncertainty. First there is the simple scenario approach, in which the valuation committee decides cash flows and/or interest rate evolution scenarios as bad, neutral and good. This is unfortunately effectively the pinnacle of valuation techniques in some investment companies.

Second type concerned with uncertainty is ENPV valuation.

3.1.2 ENPV valuation

The simplest approach how to handle uncertainty is to use expected values of the random variables that occur in the classic NPV equation. Now it is accepted that cash flow C_t and also discount rate is a random variable. In both cases it is expected that the distribution of these variables is known. The ENPV valuation is then:

$$ENPV = \sum_{t \in T} E\left[\frac{C_t}{r^t}\right] = \sum_{t \in T} \frac{p(C_t)C_t}{p(r_t)r_t^t}. \quad (3.2)$$

3.1.3 Comparing investments

There could be some part where metrics that are used to favor one investment over the other here. Comparing NPVs only seems not a good idea, but for example ROIC seems much better and the simplest strategy I would use for investment decision making.

Also comparing IRR and other statistics seems like a magic, maybe because in the end the decision is only backed by these numbers, but the feeling manager has is what the decision is based upon.

3.1.4 Other frequently used valuation techniques

Probably not needed, thus it should not be here. Just an idea. IRR, EVA, EPS,...

3.1.5 Decision tree analysis

The decision tree analysis, which was introduced in... is the first valuation technique that acknowledges the importance of further management of investments. In some sense it is equivalent to a NPV with a number of scenarios equal to number of leaves in the decision tree representing an investment.

Decision tree analysis is based on creation of a decision tree, with three type of nodes. Decision nodes serve as a representation of an opportunity to take action by the decision maker, whereas *probability* nodes serve as a representation of an uncertain event that cannot be modified.

As can be seen in many many articles the added value of acknowledging the possibility to act upon the existing investment can be crucial. In more dynamic environments this value increases, since there is generally a bigger change between cash flows of different actions.

It bewilders me, from the terminology perspective, how some authors interpret a difference between two different DTA trees as an "Real option". However from the linguistic point of view that makes sense. A new branch in a decision tree is literally a new option (real option) to act upon resolution of uncertainties in an investment.

3.2 Real option analysis

The idea to analyze investment opportunities and their parts as options that are available to a decision maker came right after a widely recognized valuation technique of financial options was adopted. This technique is of course the Nobel-prize winning Black-Scholes-Merton model, first published in [3] which allows to value financial option with only small requirements and a widely accepted assumptions.

Based on this valuation model, group led by ... came up with the idea to adopt this type of thinking into real project investment decision. The similarities between simple project investments and financial options are clear, but for more complicated projects, this interpretation starts to be less and less viable. Further discussion is left to next chapters.

3.2.1 Financial options

The rise of the derivative market of futures and options comes from the idea of insurance against the fluctuation of commodity prices on the market. To secure against the unwanted shifts in the price of a commodity viable for running a business certainly has value for the participants of the economy. The question long time (?) was how large this value is.

A financial option is an ability to buy or sell an asset for a given price in some time in the future. A financial option can be also defined by its two properties, price for which shares can be bought (in case of call option) or sold (put option) - strike price, and the time at which the option can be realized - time to maturity. Furthermore there are two types of options european and american, which have the ability to execute the option earlier than the maturity time. Because it was proved that executing an option earlier than in the time of maturity [], it does not matter which options are we going to talk about.

<Also talking only about calls since puts are not important in real options>

An example of a financial option from current market can be the option to buy Amazon stock currently valued 2286 for 3000 USD in 17-7-2020. This option has a price tag of 11 USD at the time of writing.

Option speculation is interesting because it allows the investor large profits if the markets move in the right direction. For example if the price of Amazon stock from the previous example will be 3110 USD, then an investor into those type of options can count with a 1000% profit.

The question now is, where did the price tag for an option came from. A elegant answer lies in the Black-Scholes model for option pricing.

3.2.2 Black-Scholes Merton model

Only four parameters and one assumption is needed to determine a value of an option according to BSM model for option pricing. Assume that the market is complete, and thus the law of one price holds []. Then to value a option you need to know only its time to maturity, its strike price, the current price of the underlying stock and its volatility as follows [2]:

$$C = S N(d_1) - PV(K)N(d_2), \quad (3.3)$$

where S is the strike price, $PV(K)$ is a price of a bond paying K on the expiration day of the option and $N(d)$ is a cumulative normal distribution, probability that a normally distributed variable is less than d . Value of d_1 and d_2 is then defined as:

$$d_1 = \frac{\ln(S/PV(K))}{\sigma \sqrt{T}} + \frac{\sigma \sqrt{T}}{2} d_2 = d_1 - \sigma \sqrt{T} \quad (3.4)$$

The dependency of the price of an option is positive in case of volatility and time to maturity Increasing these parameters leads to a higher option price. On the contrary the rise in current stock price or strike price of the options lowers the value of an option.

3.2.3 Real options

<Difference in DTAs, analogy for financial options, or CAPM model?>

3.3 Statistical decision theory

The second pillar upon which this thesis stands is the statistical decision theory. An area of applied mathematics with broad history. The class of problems that can be solved by this approach is very wide and greatly standardized.

The STDs main focus is to determine the optimal decision to act upon in dynamic and uncertain environment. A classical structure of a decision making problem is: ...

In SDT the uncertainty is modeled by the transition probability function which is either assumed to be known, or it is being estimated by either classical statistic methods or Bayesian methods.

The optimal decision strategy (sequence of actions) is the one that gives the maximal cumulative reward. This maximization can be total (in finite or discounted cases) or per period. We will focus on the total cumulative reward in finite processes (?)

To compute value function in all decision nodes of the grid of statistical decision theory a smart idea of backward induction, called dynamic programming is used. The narrative that a sequence of actions is optimal if and only if the last action of the sequence is clear, but powerful when used in this context. Instead of going through <something like S^{A^B} possible paths and computing expected cumulative reward for all of them, one can optimize his actions from the horizon, lowering the number of computations needed significantly to an order of <something like $S^A * N$...>. For example if the scale of the problem is ... the change is from 10^{100} to 10^{20} .

To obtain a value function in each decision node of the stochastic decision process a Bellman equation is used.

$$TODO - BELLMANEQ \quad (3.5)$$

When valuing an investment both statistical approaches to estimation are used, but in different cases. The statistical theory allows to estimate parameters when a lot of data is present (for example to determine a volatility of a stock). On the other hand, when the data is scarce and the decision making is undertaken by an expert, the Bayesian theory enables to import his subjective idea into the process.

3.3.1 Bayesian statistics

<Bayesian statics allow to include subjective opinion of the uncertainty> <Bayesian statistics are based on randomization of a parameter> <There is a model for the data in the same way that in classical statistics> <Not even classical statistics is objective, since you always have to choose a model> <Bayes equation, updating on data, certainty equivalence,...>

3.3.2 Utility

When rewards are not valued linearly by the decision maker the concept of utility comes in. One of the simplest examples to demonstrate the usage of utility is given by [1]. Imagine you are given a choice, either get 500\$ right away or gamble for 1000 \$ in a fair coin toss. A rational decision maker driven only by the expected value given should be indifferent to these choices, but the majority of people tend to take the certain option []. The decision making is clearer as the amount of money rises, there is a little difference for an average human to obtain 10M and 20M, the change in his life will be almost the same with both results. However one is certain and the other is not.

Another interesting example of utility is the St. Petersburg paradox <find reference> where rational decision maker should be willing to pay unlimited amount of money to be able to play the game, but people are seldom willing to pay more than ...

There is also an interesting asymmetry between incurring a loss and getting a profit. This can be seen from picture ... [1].

Each person has their own utility function with respect to money and it can be sketched essentially by a questionnaire. More questions mean more precision, and it also keeps your answers coherent.

3.4 Approximate dynamic programming

<Why it exists and what are the types?> Actual application of classic dynamic programming

Chapter 4

Real option analysis in terminology of stochastic decision control

The similarity of ROA and SDT is clear (?).
MAIN PART OF THE THESIS

Chapter 5

Application of the cheery-picked valuation technique.

The class of valuation problems is ideal for demonstration of the usability of the newly developed valuation technique. Due to the inherit uncertainty in this field and many actions that can be undertaken, the additional value assigned to the investment opportunity can be significant.

We start with a rigorous mathematical definition of this class of valuation techniques in terms of SDT. Then we pick one example and promptly show that the number of possible states is exponential with respect to... Approximate dynamic programming techniques like Q-learning or SARSA were developed in order to solve exactly these types of problems. Due to their strengths and only a minor flaws their usage is justified.

5.1 Approximate dynamic programming

<Proposal of a fitting method to the class of problems defined above>

5.2 Valuation example

5.2.1 NPV valuation

5.2.2 DTA valuation

5.2.3 Classic ROA valuation

5.2.4 New approach

Chapter 6

Discussion

The new approach is better because it solves the current problems with ... Also the applicability of the new approach is in my opinion broader since it can address multiple sources of uncertainty. Furthermore the power of the decision making process is kept in the hands of the decision maker through creation of prior distributions. The manager is guided through the world of utility functions and priors, which both can be created from a set of simple questions about gambles and beliefs of the manager. The creation and usage of the utility and prior density functions are fool-proof in a sense of mathematical coherence.

Chapter 7

Conclusions

In my master's thesis I have rigorously compared the state-of-the-art valuation techniques used in present investment companies. I have shown the advantages and disadvantages of real option analysis and stochastic decision theory. The combination of these, which I call ..., yields a new view on the world of risky investments that empowers the decision maker and thus allows for better adoption in the rigid environment of investing.

Chapter 8

Bachelor's thesis parts that could be useful for TeX styling

8.1 Discrete Markov decisions processes

A lot of real life decision tasks can be approximately described by the mathematical framework called discrete Markov decision processes (MDPs), Fig. XY. The definition of this mathematical framework is crucial for this work, as it is the basis upon which the theory of policy optimisation is built.

Definition 8.1. *Discrete Markov Decision Process \mathcal{M} is defined as ordered set of five elements $(\mathbf{T}, \mathbf{S}, \mathbf{A}, P, R)$, where:*

- \mathbf{T} stands for a discrete, finite set of decision epochs; $\mathbf{T} = \{0, 1, 2, \dots, |\mathbf{T}|\}$, where¹ $|\mathbf{T}| \in \mathbb{N}$
The state in time epoch $t = 0$ is known and it is denoted s_0 .

Definition 8.2. *Let $\mathcal{M} = (\mathbf{T}, \mathbf{S}, \mathbf{A}, P, R)$ be a MDP and $\pi_t \in \Pi_t$ a policy, then **value functions** of \mathcal{M} , $\varphi_t^{\pi_t} : \mathbf{S} \rightarrow \mathbb{R}$ are defined $\forall \pi_t \in \Pi_t, \forall t \in \mathbf{T} \setminus \{N\}$ as:*

$$\varphi_t^{\pi_t}(s) = E \left[\sum_{\tau > t, \tau \in \mathbf{T}} R(s_\tau, a_{\tau-1}, s_{\tau-1}) | s_t = s \right], \quad (8.1)$$

Theorem 1. *Let $\mathcal{M} = (\mathbf{T}, \mathbf{S}, \mathbf{A}, P, R)$ be a MDP, then the optimal value function $\varphi_t^o(s) = \max_{\pi_t \in \Pi_t} \varphi_t^{\pi_t}$ can be computed recursively $\forall t \in \mathbf{T} \setminus \{N\}$ through the equation:*

$$\varphi_t^o(s) = \max_{a_t \in \mathbf{A}} E \left[R(s_{t+1}, a, s) + \varphi_{t+1}^o(s_{t+1}) | a_t, s \right], \quad (8.2)$$

considering that $\varphi_N^o = 0$ as Definition 8.2 implies. Furthermore the optimal policy π_t^o , as the argument of maxima, is concentrated on maximising arguments in Equation (8.2).

Proof. This theorem will be proven via a finite backward mathematical induction.

At first, let us take an action a_{N-1}^o in the time epoch $N - 1$ defined as:

$$a_{N-1}^o(s) = \text{Arg max}_{a \in \mathbf{A}} E[R(s_N, a, s) | a, s]. \quad (8.3)$$

By the definition of maxima the inequality

□

Algorithm 1 Finding the optimal policy for a *single system* MDP with known P

Require: $\mathcal{M} = (\mathbf{T}, \mathbf{S}, \mathbf{A}, P, R)$

- 1: $\varphi_N^o(s) \leftarrow 0, \forall s \in \mathbf{S}$ ▷ Based on Definition 8.2.
 - 2: $t \leftarrow N$
 - 3: **while** $t \neq 0$ **do**
 - 4: **for** each $s \in \{1, 2, \dots, |\mathbf{S}|\}$ **do**
 - 5: $\varphi_{t-1}^o(s) \leftarrow \text{Equation (8.2)}$ ▷ With known P and φ_t^o
 - 6: $t \leftarrow t - 1$
 - 7: $\pi_0^o(s) \leftarrow \operatorname{argmax} \varphi_0^o(s) \forall s \in \mathbf{S}$, ▷ Deriving the optimal policy
 - 8: **return** π_0^o
-

We can now describe the dynamic programming algorithm in detail.

Example 1. A real life example of this MDP could be again a biased coin tossing. Imagine a game when you are sitting at the table on which there are multiple coins with generally different bias. You get one coin on certain side in your hands and you are told the rules. You get \$1 each time the coin lands on Heads. You can choose to change your coin for any other coin for \$3 and you have to place it to the side it is on the table (you have to "conserve its state"). The turn of the coin costs 5 cents and the result of the toss generally depends on the side from which the coin is tossed. You know the bias of each one coin and the question is: what is the best sequence of actions that you can perform to get as much money as possible, on average, from this N -round game?

It is important to mention that you are able to see the other coins "on the table" so that the position state s^p is known in each time epoch.

Discussion The main goal of this simulation is to check the quality of the experimentally derived sub-optimal policy via Algorithm 1. As we can see, the difference between the average cumulative reward of the derived policy and the expected optimal average cumulative reward is approximately 0.5% in the first simulation with 500 Monte Carlo iterations and approximately 0.026% in the second simulation with 5000 iterations.²

We can also see that in the second performed experiment made for 5 000 iterations, ten times the first case, both the variance of the distribution and the difference from the optimal value decreased. It is then assumed that this trend would continue with more iterations and that both the variance and the difference would get closer and closer to zero.

With this assumption we can say that the experimentally derived optimal policy is in fact the optimal one.³

¹Because the number of time epochs $|\mathbf{T}|$ is used frequently, the notation $|\mathbf{T}| = N$ is used.

²I dont know how to fit the exceedance in here, so if it is not that important I would prefer not to talk about it.

³Is that a good reasoning?

$P(:, :, s = 1)$		
\tilde{s}/a	1	2
1	0.4237	0.6885
2	0.5763	0.3115

$R(:, :, s = 1)$		
\tilde{s}/a	1	2
1	1.4315	-0.0569
2	2.7287	-0.3111

$w(:, :, s = 1)$		
\tilde{s}/a	1	2
1	8	5
2	1	2

$P(:, :, s = 2)$		
\tilde{s}/a	1	2
1	0.7282	0.8816
2	0.2718	0.1184

$R(:, :, s = 2)$		
\tilde{s}/a	1	2
1	-0.1039	2.5921
2	1.1251	-0.2876

$w(:, :, s = 2)$		
\tilde{s}/a	1	2
1	1	2
2	9	1

Table 8.1: Tables of P , R and w values for the CE approach simulation. Original 3D arrays are presented by two "slices" which differ in the initial state $s = 1$ or $s = 2$.

⁴Hope this representation is OK, and how to do the backslash \tilde{s}, a ?

Bibliography

- [1] S. Bacci and B. Chiandotto. *Introduction to Statistical Decision Theory: Utility Theory and Causal Analysis*. CRC Press, 2019.
- [2] J. Berk and P. DeMarzo. *Corporate Finance*. Pearson, 2014.
- [3] Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.
- [4] T.E. Copeland and V. Antikarov. *Real Options: A Practitioner's Guide*. Texere, 2001.
- [5] G. Guthrie. *Real Options in Theory and Practice*. Financial Management Association Survey and Synthesis. Oxford University Press, 2009.
- [6] A. Vollert. *A Stochastic Control Framework for Real Options in Strategic Evaluation*. Birkhäuser Boston, 2012.