Portfolio Theories

Philippe J.S. De Brouwer*

September 1, 2010

Abstract

A brief introduction to the contemporary literature about portfolio theories.

Contents

1 Introduction 3

2 Chronological Overview 3

2.1 Homo Rationalis: XVII - XIXth Century . . . . . . . . . . . . 4
  2.1.1 Expected Value . . . . . . . . . . . . . . . . . . . . . . 4
  2.1.2 The Law of Large Numbers . . . . . . . . . . . . . . . 5
  2.1.3 St Petersburg Paradox . . . . . . . . . . . . . . . . . . 5
  2.1.4 Expected Utility Theory (EUT) . . . . . . . . . . . . 5
  2.1.5 Cramer’s Solution . . . . . . . . . . . . . . . . . . . . 6

2.2 Behavioural Evidence and Thinking: 1750 till 1950 . . . . . 7
  2.2.1 Adam Smith (1759) . . . . . . . . . . . . . . . . . . . 7
  2.2.2 Early mentioning of irrational behaviour: Mackay (1841)
    and others . . . . . . . . . . . . . . . . . . . . . . . . . 8
  2.2.3 The Ellsberg Paradox described by John F.M. Keynes
    in 1921 . . . . . . . . . . . . . . . . . . . . . . . . . . . 9
  2.2.4 John F.M. Keynes and Chaos theory (1936) . . . . . . 10
  2.2.5 Maslow’s theory of Hierarchy of Needs (1943) and the
    framing heuristic . . . . . . . . . . . . . . . . . . . . . . 19
  2.2.6 The Axioms of Von Neuman and Morgenstern . . . . . 25
  2.2.7 The Friedman-Savage Puzzle (1948) . . . . . . . . . . 27

2.3 Rational Portfolio Theories (1950s and 1960s) . . . . . . . . 29
  2.3.1 Modern Portfolio Theory (1952) . . . . . . . . . . . . 29

*Philippe J.S. De Brouwer is Executive Director at Eperon Asset Management Ltd.
(fully owned subsidiary of KBC Group N.V. – The views expressed by the author are
strictly his, and do not necessarily represent those of KBC Group N.V nor any of its
subsidiaries. – email: philippe@de-brouwer.net
2.3.2 Markowitz' Customary Wealth Theory (1952) ..... 34
2.3.3 Roy's Safety First Portfolio Theory (1952) ..... 38
2.3.4 The Allais Paradox (1953) ......................... 40
2.3.5 Subjective Expected Utility Theory (SEU) (1954) .. 41
2.3.6 cognitive dissonance (1956) ....................... 42
2.3.7 The Ellsberg Paradox popularized by Daniel Ellsberg (1961) ........................................ 42
2.3.8 The Capital Asset Pricing Model (CAPM – 1961) .. 42
2.3.9 The Fallacy of Large Numbers (1963) ............. 49
2.3.10 The Efficient Market Hypothesis (EMH) (1964) ... 49
2.3.11 Models for Lifetime Portfolio Selection - 1968 and 1969 51
2.3.12 An Intertemporal Capital Asset Pricing Model - 1973 53
2.4 Upcoming challengers from Psychology: the 1970s and the early 1980s ........................................ 53
2.4.1 Availability Heuristic (1973) ....................... 53
2.4.2 More Heuristics (1974) ............................. 54
2.4.3 Support for CAPM and related theories – 1976 .... 54
2.4.4 Prospect Theory (1979) ............................. 55
2.4.5 Framing (1981) ....................................... 55
2.4.6 The Volatility Puzzle (1981) ....................... 56
2.4.7 Judgemental Biases Described (1982) ............... 56
2.5 First Signs of Acceptance for Behavioural Finance: 1985 ... 57
2.5.1 The first real evidence of market non-efficiencies ... 57
2.5.2 Mental Accounting .................................. 57
2.5.3 The Equity Premium Puzzle ....................... 57
2.6 Towards Acceptance of Behavioural Finance Among Scholars (1985 - 2000) .................................... 58
2.6.1 The late 1980s ...................................... 58
2.6.2 Evidence of Analyst overreaction .................. 61
2.6.3 The Discussion about the Efficiency of the Stock Markets ................................................. 61
2.6.4 Loss Aversion and endowment effect ............... 62
2.6.5 And more Contributions to the Behavioural Paradigm 62
2.6.6 Cumulative Prospect Theory (1992) ............... 62
2.6.7 A Solution for the Equity Premium Puzzle (1995) .. 63
2.6.8 Vivid Interest in Utility Optimizing Strategies – 1995-1997 .................................................. 65
2.6.9 More Evidence ...................................... 66
2.6.10 Critics to Behavioural finance (1998) ............. 67
2.6.11 An avalanche of evidence, theorems and books (1998 - 2000) ............................................. 67
2.6.12 Behavioural Portfolio Theory (BPT – 2000) ........ 69
2.6.13 Fallacy of Large Numbers Revisited (2001) ...... 71
2.7 2002: an Excellent Behavioural Finance Year ........... 72
1 Introduction

It is worth to build foundations solidly before starting any further construction. In this chapter we illustrate the scientific thinking that paved the road to Maslowian Portfolio Theory (MaPT) and Target Oriented Investment Advice (TOIA).

We try to find the original ideas and original authors of the different theses in order to illustrate the logic in the scientific development and thinking about portfolio theories. Therefore we emphasize not only new theories, but also the paradoxes or puzzles that were a direct incentive to produce this new theory.

2 Chronological Overview

During the last century two paradigms competed for general acceptance:

“The Rational Paradigm” here investors are rational and optimize their utility function (which is smooth, concave and based on final wealth) in order to make decisions, markets are efficient, and each investor needs one optimal portfolio;

“The Behavioural Paradigm” here investors’ behaviour displays important biases compared to the rational behaviour (for example, people are loss averse, overconfident, and think in “frames”), utility is relative to a reference point and can display concave and convex areas, and therefore investors do not have one efficient portfolio but have fragmented portfolios, markets are not efficient (and display for example trends and mean reverting patterns).

This struggle between the two paradigms was most pronounced in the XXth century. Success shifted between the paradigms during that century multiple times, and the most prominent economists and psychologists were involved. Now it seems widely accepted that investors do display important deviations from rational behaviour. It is also a fact that patterns with an
information content bigger than zero have been detected, and even out of sample studies show on past data (with hindsight) the possibility to outperform some markets.

It is however not so straightforward do claim that this also means the end of the EMH. Even the availability of studies (such example see (Caginalp and Laurent 1998)) where on past data statistically significant profitable strategies can be found, (to the author’s best knowledge) we lack successful applications of such theory. Even if a mechanism worked for the last decades, there is no guarantee that it will to continue for the next decade.\(^1\)

In order to give insight in the evolution of the thinking about investment portfolio selection, we present the developments chronological. One exception is the very short introduction to chaos theory, as it is rather independent from the rest of this work\(^2\) (it is all grouped in one chapter).

2.1 Homo Rationalis: XVII - XIX\(^{th}\) Century

2.1.1 Expected Value

Blaise Pascal (June 19, 1623 – August 19, 1662) was challenged by a friend, Antoine Gombaud (self acclaimed “Chevalier de Méré” and writer), with a gambling puzzle. The brain teaser was that two players who want to finish a game early and, given the current circumstances of it, want to divide the stakes fairly, based on the chance each has of winning the game from that point. How should they find this “fair amount”?

In 1654, he corresponded with Louis de Fermat on the subject of gambling. And it is in the discussion about this problem, that the foundations of the mathematical theory of probabilities are laid. From this discussion, the notion of expected value was introduced.

de Fermat and Pascal are generally considered as the founders of the “theory of probability”, and their work laid the foundations for Leibniz’s formulation of the infinitesimal calculus.\(^3\)

Pascal used later (in the posthumous published “Pensées”) a probabilistic argument, Pascal’s Wager, to justify belief in God and a virtuous life.\(^4\)

---

\(^1\)Mathematically chaotic time series might display patterns for a while, but they can change as time flow. More insight in this phenomenon is given in Chapter 2.2.4 on page 10

\(^2\)This is done, because this brief introduction to chaos theory provides important insight in the nature and the behaviour of financial markets. This insight is has many applications in financial theories and behavioural finances as developed later in this work.

\(^3\)Actually, de Fermat was the first person known to have evaluated the integral of general power functions. Using an ingenious trick, he was able to reduce this evaluation to the sum of geometric series. The resulting formula was helpful to Newton, and then Leibniz, when they independently developed the “fundamental theorem of calculus” (that relates differentiation and integration), that was in its turn the base of and a necessity for the development of sciences for the centuries.

\(^4\)Pascal’s Wager (or Pascal’s Gambit) is a suggestion posed by the French philosopher Blaise Pascal that even though the existence of God cannot be determined through reason,
In the next century, Daniel Bernoulli mentions that expected value is used
“ever since mathematicians first began to study the measurement
of risk”
– (Bernoulli 1738) - English translation in (Bernoulli 1954).

2.1.2 The Law of Large Numbers

The Italian mathematician Gerolamo Cardano (1501 – 1576) stated -without proving it- that the accuracy of statistical experiments tends to improve with the number of times that the experiment is repeated (Mlodinow 2008). This idea was later formalized as a law of large numbers (LLN henceforth). The LLN was first proved by Jacob Bernoulli. It took him over 20 years to develop a sufficiently rigorous mathematical proof which was published in his Ars Conjectandi (Bernoulli 1713a) – chapter 4. He named this his “Golden Theorem” but it became generally known as “Bernoulli’s Theorem”. This should not be confused with the principle in physics with the same name, named after Jacob Bernoulli’s nephew Daniel Bernoulli. It is reported that we owe the name “La loi des grands nombres” (“The law of large numbers”) to Siméon Denis Poisson.

2.1.3 St Petersburg Paradox

Nicolas Bernoulli first stated this paradox in a letter to Pierre Raymond de Montmort of 9 September 1713 (Bernoulli 1713b).

Imagine the following game: I toss a coin, until it lands “head”. Every
time that I toss, I will pay you: $2^{N-1}$€ ($N$ is the number of the toss). The
game ends when the coin lands “tail”.

How much would you willingly pay in order to participate in such game?

The expected return of this game is infinite, but no reasonable person
would give all his belongings in order to be allowed play this game.

The name of the paradox is based on Daniel Bernoulli’s presentation of
the problem and his solution, published in 1738 in the Commentaries of the
Imperial Academy of Science of Saint Petersburg (Bernoulli 1738). However,
the problem was invented by Daniel’s cousin Nicolas Bernoulli as mentioned
earlier.

2.1.4 Expected Utility Theory (EUT)

The utility function was introduced by Gabriel Cramer, when in 1728 in
order to solve the St Petersburg Paradox (Cramer 1728). He wrote:

a person should “wager” as though God exists, because so living has everything to gain,
and nothing to lose. It was set out in note 233 of his “Pensées”, a posthumously published
collection of notes made by Pascal in his last years as he worked on a treatise on Christian
apologetics.
“(...) in their theory, mathematicians evaluate money in proportion to its quantity while, in practice, people with common sense evaluate money in proportion to the utility they can obtain from it.”
(Cramer 1728)

The Expected Utility Theory involves the explicit use of a utility function, an expected utility hypothesis, and the presumption of diminishing marginal utility of money.

Ten years later Daniel Bernoulli wrote in his 1738 landmark article:

“The determination of the value of an item must not be based on the price, but rather on the utility it yields. There is no doubt that a gain of one thousand ducats is more significant to the pauper than to a rich man though both gain the same amount.”
(Bernoulli 1738) - translated to English in 1954.

A common utility model, suggested by Bernoulli himself, is the logarithmic function

\[ U(W) = \ln(W) \]  

(known as “log utility”). It is a function of the gambler’s total wealth \( W \), and the concept of diminishing marginal utility of money is built into it. Under the expected utility hypothesis, expected utilities can be calculated the same way expected values are. For each possible event, the change in utility

\[ \ln(\text{wealth after the event}) - \ln(\text{wealth before the event}) \]  

will be weighted by the probability of that event occurring.

Let \( C \) be the cost charged to enter the game. The expected utility of the lottery now converges to a finite value:

\[ EU = \sum_{k=1}^{\infty} \frac{\ln(W + 2^k - C) - \ln(W)}{2^k} \]

\[ < \infty \]  

2.1.5 Cramer’s Solution

The log utility will work (see Equation 3 on this page) for the exact pay-off proposed in the original formulation of the St Petersburg paradox. However it appears that for each non bounded utility function one can define a pay-off that makes the expected utility infinite. For example we could propose a pay-off of \( e^{2N^2} C \) for each toss, and the log-utility will again diverge to infinity for the St. Peterburgs Paradox.
Cramer proposed therefore that a utility function should be bound by a maximum value. Unbound utility functions can be a good approximation when one is only interested in the phenomenology for finite values. However, the simple observation that when one owns all money on earth, that this is the moment that money has no value any more; the same holds for unreasonably high quantities of any good.\footnote{One bread a day is very valuable, when one gets ten breads it becomes possible to trade ... but what would one do with a googol ($=10^{100}$) of breads? Not only they are useless, but they their shear presence would become a serious burden. So actually for very large numbers of goods not only $U(\infty) \leq a$, but also $U(\infty) = -\infty$.} The fact that a utility function is necessary bounded was first mentioned by Karl Menger (mathematician and son to Carl Menger who was an economist) in (Menger 1934) and is sometimes referred to as “Karl Menger Paradox”.

2.2 Behavioural Evidence and Thinking: 1750 till 1950

2.2.1 Adam Smith (1759)

Adam Smith’s “The Wealth of Nations” (Smith 1776) helped to create the discipline of economics with its conjuring of the invisible hand, self-interest, and other explanations of market forces that have influenced academics, governments, and business leaders ever since. However in Adam Smith’s earlier work “The Theory of Moral Sentiments” (Smith 1759), that we find him as a behavioural economist. Providing ideas that lead directly to loss aversion, overconfidence, myopic loss aversion, and some other ideas.\footnote{More details can be found for example in (Nava Ashraf and Loewenstein 2006). The authors find that Smith’s insights from 1759 can contribute to modern thinking on everything from our fascination with celebrity to the theory of loss aversion.}

Adam Smith presents economic actors in “The Theory of Moral Sentiments” as people driven by an internal struggle between their impulsiveness, passions, and their “impartial spectator”. The passions include drives such as hunger, sex, and emotions such as fear and anger. Smith viewed behaviour as under the direct control of the passions, but believed that people could override passion-driven behaviour by viewing their own behaviour from the perspective of an outsider, “the impartial spectator”: like a person, looking over the shoulder of the economic actor, that scrutinizes every move he makes.

Further Adam Smith provides interesting examples of non rational behaviour. People weigh out-of-pocket costs more than opportunity costs, have self-control problems and are overconfident. They display erratic patterns of sympathy, but are consistently concerned about fairness and justice. They are motivated more by ego than by any kind of direct pleasure from consumption and -though they don’t anticipate it-, ultimately derive little pleasure from either.

In short, in Adam Smith’s thinking, the world is not inhabited by purely...
rational and self-interested actors, but rather by behavioural human beings, whose behaviour is governed by complex motivations.

### 2.2.2 Early mentioning of irrational behaviour: Mackay (1841) and others

Already in the nineteenth century, Charles Mackay described irrational behaviour (Mackay 1841) in his book “Extraordinary Popular Delusions and the Madness of Crowds”. He described for example the “South Sea Bubble” and the “Tulipomania” where the parallel with the financial markets of today is very clear. Other examples like the magnetizers, the witch mania, the alchemists, etc. at least prove that human behaviour is not always rational.

le Bon had a more psychological approach when he wrote “The Crowd: A Study of the Popular Mind” (le Bon 1896).

Selden believed that

> “movements of prices on the exchanges are dependent to a very considerable degree on the mental attitude of the investing and trading public”

– (Selden 1912)

### Early Development of the EMH

Louis Bachelier developed in 1900 already the EMH in his book “Théorie de la spécula­tion” (Bachelier 1900a) and (Bachelier 1900b). However his work remained largely unknown till P.A. Samuelson brought it to the attention in the early 1960s (after Jimmie Savage’s postcard reminding him of Bacheliers work). Afterwards E. Fama would publish his ideas in (Fama 1965), which is now the generally considered as the creator of the EMH. We will discuss this idea also later in Chapter 2.3.10 on page 49.

Actually Bachelier did much more than writing down the efficient market hypothesis. This was for him just a step in the process of creating an option pricing theory. Actually the option pricing model of Bachelier was exactly the same as what later was used by the Nobel Prize-winning solution of the option pricing problem by Fischer Black, Myron Scholes and Robert Merton in 1973. Only Bachelier did not have the mathematical tools in place to write the exact solution.

Bacheliers achievement in his thesis was to introduce, starting from scratch many aspects of stochastics: he defined Brownian motion and the Markov property\(^7\), derived the ChapmanKolmogorov equation and established the connection between Brownian motion and the heat equation.

---

\(^7\)A stochastical system has the Markov Property if the future state does not depends on the past but only on the actual state. A Markow-chain is hence a random process where the next state depends only on the previous one. What more do we need for financial markets to be efficient.
Many of those works are now generally association with other names and with significantly later dates.

2.2.3 The Ellsberg Paradox described by John F.M. Keynes in 1921

The Ellsberg paradox was brought to the general attention in 1961 by Daniel Ellsberg (Ellsberg 1961), but was earlier described by John F.M. Keynes (Keynes 1921) - pp. 75–76, p. 315, ft. 2.

The paradox demonstrates how human decision making defies the expected utility hypothesis, and it is explained by “ambiguity aversion”.

Suppose you have an urn containing 30 red balls and 60 other balls that are either black or yellow. You don’t know how many black or yellow balls there are, but that the total number of black balls plus the total number of yellow equals 60. The balls are well mixed so that each individual ball is as likely to be drawn as any other. You are now given a choice between two gambles:

- **Gamble A**: You receive € 100 if you draw a red ball
- **Gamble B**: You receive € 100 if you draw a black ball

Also you are given the choice between these two gambles (about a different draw from the same urn, after all balls have been put back into it):

- **Gamble C**: You receive € 100 if you draw a red or yellow ball
- **Gamble D**: You receive € 100 if you draw a black or yellow ball

Since the prizes are exactly the same, it follows that you will prefer Gamble A to Gamble B if, and only if, you believe that drawing a red ball is more likely than drawing a black ball (according to expected utility theory). Also, there would be no clear preference between the choices if you thought that a red ball was as likely as a black ball. Similarly it follows that you will prefer Gamble C to Gamble D if, and only if, you believe that drawing a red or yellow ball is more likely than drawing a black or yellow ball. If drawing a red ball is more likely than drawing a black ball, then drawing a red or yellow ball is also more likely than drawing a black or yellow ball. So, supposing you prefer Gamble A to Gamble B, it follows that you will also prefer Gamble C to Gamble D. And, supposing instead that you prefer Gamble D to Gamble C, it follows that you will also prefer Gamble B to Gamble A.

When surveyed, however, most people strictly prefer Gamble A to Gamble B and Gamble D to Gamble C. Therefore, some assumptions of the expected utility theory are violated.

Mathematically, your estimated probabilities of each coloured ball to be drawn can be represented as: \( R, Y, \) and \( B \) (with \( R = P[\text{the ball drawn is Red}] \),...
and similar for the two other colors). If you strictly prefer Gamble A to Gamble B, by utility theory, it is presumed this preference is reflected by the expected utilities of the two gambles: specifically, it must be the case that

\[ R \cdot U(100) + (1 - R) \cdot U(0) > B \cdot U(100) + (1 - B) \cdot U(0) \]  

(5)

where \( U(\cdot) \) is your utility function. If \( U(100) > U(0) \) (you strictly prefer €100 to nothing), this simplifies to:

\[ R > B \]  

(6)

If you also strictly prefer Gamble D to Gamble C, the following inequality is similarly obtained:

\[ B \cdot U(100) + Y \cdot U(100) + R \cdot U(0) > R \cdot U(100) + Y \cdot U(100) + B \cdot U(0) \]  

(7)

This simplifies to:

\[ B > R \]  

(8)

Obviously, 6 is in contradiction with 8. This indicates that human preferences are inconsistent with expected-utility theory.

2.2.4 John F.M. Keynes and Chaos theory (1936)

In the beginning of the twentieth century J.M. Keynes laid the basis of modern economic theory. In his landmark book “The General Theory of Employment, Interest, and Money” (Keynes 1936) he put forward that financial markets could be compared with a special beauty contest. The contest goes as follows. There is a selection of beautiful girls. You can only win the game by selecting the girl that will get the highest score when the scores of all jury members are added. In this game, you will not select the girl that according to your standards is the most beautiful, but you will select the one you think the others will like most. This is an entirely other game, with entirely different dynamics, than a game where everyone makes decisions based upon his own judgement.

The parallel with financial markets is obvious: it is of no use to select a company in which you believe. You should aim to buy shares of a company in which others believe, because if you buy and then everyone else sells, it is a disaster. A closed game where you select stocks from which you anticipate that others anticipate that others will relatively prefer them is a complex, non-linear feedback system. Mathematicians and physicians have thoroughly studied these systems since the late sixties.
**Complex Non-linear Systems.** These systems display very specific behaviour that got understood in the 1970s (see for example (Haken 1977), (Prigogine 1980), (Prigogine and Stengers 1984), and applied on the financial markets (Peters 1999) and (Trippi 1995)).

This theory was first developed for large systems far from thermodynamic equilibrium. Soon Chaos theory found its application in the most diverse sciences and systems: from the prediction of weather over a dripping faucet to biological evolution.

Till then scientists were mainly focused on thermodynamic systems close to equilibrium, the behaviour of such systems is linear. Those systems could easily be integrated and for centuries studying their behaviour was interesting research. A simple example of such system is heat transfer in liquid and solid bodies. The equation describing heat transfer is in a first approximation: \( \frac{dT(t)}{dt} = -a(T - T_{env}) \). If there is a temperature gradient, the temperature will adapt and cool down or heat up proportional (linear) to the difference in temperature. This equation is easily integrated, and it’s solution is: \( T(t) = T_{env} + (T(0) - T_{env}) e^{-at} \).

This equation states that the temperature is predictable, and moving smoothly. A similar result also holds for the heat transport in a liquid. For example when one starts heating water from below, the heat moves up at a speed proportional to the difference in temperature with the air above. For each small cube of water inside the tank, the temperature can exactly be described as above. This is linear behaviour.

But when one increases the heat source (and hence the temperature

---

8This chapter is a very brief introduction to chaos theory, however it provides important insight in the nature and the behaviour of financial markets. A more profound introduction for the financier is “Chaos & non-linear Dynamics in the Financial Markets” (Trippi 1995). This book explains the theory, provides evidence of chaos in financial markets, and even present possible applications for practitioners. A more accessible work is the book “Patterns in the Dark” (Peters 1999).

For an example of complex patterns (even 2 dimensional) and how they appear in nematic liquid crystals we point with pleasure to (De Brouwer and Walgraef 1993), actually the whole book “Instabilities and Nonequilibrium Structures IV” (Tirapegui and Zeller 1993), in which the article is published, provides many examples.

9A system is said to be “large” or “complex” when it consists of so many sub-systems, that the system cannot be understood by studying the interactions of the sub-systems. A macroscopic view is neccessary. For example the individual water molecules will not really help us to describe how water behaves, we need a macroscopic approach such as the Navier-Stokes equations that threat water as a continuum, in stead of a N-molecule system.

10Biological evolution is in contradiction with the second law of thermodynamics because in biological evolution displays transitions to more complex states. These more complex states display less and less entropy. Which is in contradiction with the second law of thermodynamics that states that entropy should increase in any closed system: \( \frac{dE}{dt} \geq 0 \). Non-linear feedback systems do however show transitions to more complex states - which is essential for biological evolution. The key here is that the second law of thermodynamics does not hold for open systems (only for closed systems), and the earth is not a closed system.
gradient), at a certain point something dramatically happens. The water cannot transport the heat efficient enough by diffusion alone, and currents appear in the water. That is totally outside the description of the linear system: that is the behaviour of a non-linear system. Where any small (virtual) cube of water inside the tank was first not moving at all, now there are currents at a scale that is billions of times larger than the scale of the molecules. Symmetry has suddenly been broken, and information has been created: at some points the water is moving up, at others the water is moving down. Our linear model cannot explain this behaviour. A simple generalization to a simple non-linear model might already show the behaviour that we just described.

\[ \frac{dT(t)}{dt} = -a[(T - T_{\text{env}}) - (T - T_{\text{env}})^2] \]

In the remainder of this chapter, we will study the behaviour of similar systems and show indeed that they are able to describe behaviour that was just described. This behaviour of non-linear systems far from thermodynamic equilibrium is in many aspects very different from linear systems:

- Sudden transitions to other states (bifurcations), multiple stable states, generally accompanied with a breaking of symmetry.\(^\text{11}\)
- Eventually transition to chaos. This chaotic state is very different from a pure random state, and it is characterised by:
  - critical dependence on certain initial parameters: when a starting condition differs only very small from another, the system will sooner or later behave completely different in both cases.
  - and therefore it is impossible to predict much at long term: indeed in order to predict something, we would have to know the starting parameters with an infinite precision.

**A Simple Pricing Model: the Logistic Map.** Let us have a short look at another very simple example: we will construct a model for the price \( P \) of a certain stock. The first simple approximation could be that investors are attracted by the success of the company (measured by the price of the company). Linear systems on the contrary would display one type of behaviour, i.e. there would be one equilibrium state. The system would move gradually to another level of the same state, certainly no sudden changes. A simple example could be the heating of water from below. Till a certain heat gradient, the system behaves linearly and there is a continuous dissipation of heat from below, the higher the gradient the faster the heat transport. That is a linear system. But at a certain moment, the water itself starts to move (not only the heat), this is a sudden change, and a breaking of symmetry, a bifurcation (each water molecule will have to choose up or down). Interesting is also that now patterns are formed billions of times larger than the constituting molecules.

\(^\text{11}\)Linear systems on the contrary would display one type of behaviour, i.e. there would be one equilibrium state. The system would move gradually to another level of the same state, certainly no sudden changes. A simple example could be the heating of water from below. Till a certain heat gradient, the system behaves linearly and there is a continuous dissipation of heat from below, the higher the gradient the faster the heat transport. That is a linear system. But at a certain moment, the water itself starts to move (not only the heat), this is a sudden change, and a breaking of symmetry, a bifurcation (each water molecule will have to choose up or down). Interesting is also that now patterns are formed billions of times larger than the constituting molecules.
stock itself). Hence the price dynamics would be described by $P_{t+1} = a P_t$.

In our simple model, the price would drop to zero for $a \in [0, 1]$, be constant for $a = 1$, and explode for $a > 1$.

This linear system has one stable solution depending on the exact value of the growth parameter $a$: $0, 1, \infty$. Never there are multiple stable solutions for one growth parameter, and the original symmetries are preserved.

This model does not capture much interesting behaviour, it might be too simple. We can write a non-linear generalization of our first simple model.

Assume now that the price $P$ of a certain stock is driven by the following simple dynamics. The price in period $(t + 1)$ will depend on the price in the period $t$. More precise, the price will increase with $(a - 1)\%$ for sufficiently small values $P_t$, however if the price is rising too high, a downwards pressure will set in and contribute negatively. This downwards pressure then equals $-a P_t^2$, and becomes (both relatively and absolutely) stronger the more $P_t$ is higher.

These dynamics are described by the following equation.

$$P_{t+1} = a(1 - P_t)P_t$$

This mapping of $\mathbb{R} \mapsto \mathbb{R}$ is known as the “logistic map”.\(^{12}\) It is a polynomial mapping of degree 2, and it shows very well how simple (but non-linear) dynamics can lead to complex, chaotic behaviour.

---

\(^{12}\)The logistic map was popularized in a seminal paper by the biologist Robert May (May 1976), in part as a discrete-time demographic model analogous to the logistic equation. This equation was created by Pierre François Verhulst. Pierre François Verhulst (October 28, 1804, Brussels, Belgium – February 15, 1849, Brussels, Belgium) was a mathematician and a doctor in number theory from the University of Ghent in 1825. Verhulst published in 1838 the logistic equation (Verhulst 1838) as a model for population development:

$$\frac{dN}{dt} = aN \left(1 - \frac{N}{K}\right)$$

where $N(t)$ represents number of individuals at time $t$, $a$ the intrinsic growth rate and $K$ is the carrying capacity, or the maximum number of individuals that the environment can support. This model was rediscovered in 1920 by Raymond Pearl and Lowell Reed, who promoted its wide and indiscriminate use. The logistic equation can be integrated exactly, and has solution

$$N(t) = N(0) \frac{e^{-at}}{1 + CKe^{-at}}$$

where $C = 1/N(0) - 1/K$ is determined by the initial condition $N(0)$. The solution can also be written as a weighted harmonic mean of the initial condition and the carrying capacity.

$$\frac{1}{N(t)} = \frac{1}{K} + \frac{e^{-at}}{N(0)}$$

Although the continuous-time logistic equation is often compared to the logistic map because of similarity of form, it is actually more closely related to the Beverton-Holt model.
The continuous form of equation 9 is known as the “logistic equation”

$$\frac{dP}{dt} = aP(1 - P)$$  \hspace{1cm} \text{(logistic equation)}

and its solution

$$P(t) = \frac{1}{1 + \left(\frac{1}{P_0} - 1\right)e^{-at}}$$ \hspace{1cm} \text{(sigmoid function)}

is known as the “logistic function” or as the “sigmoid function”. The logistic function is known as the solution of the simple first-order non-linear differential equation $\frac{dP}{dt} = aP(1 - P)$ where $P$ is a variable with respect to time $t$ and with boundary condition $P(0) = P_0$. This equation is the continuous version of the logistic map. One verify the solution to be $P(t) = \frac{e^{at} + \frac{1}{e^{at} - 1}}{1 + \frac{1}{P_0} - 1}$.

Choosing the constant of integration so that $P_0 = -1/2$ gives the well-known form of the definition of the logistic curve $P(t) = \frac{e^{at}}{1 + e^{at}} = \frac{1}{1 + e^{-t}}$.

The logistic function is the inverse of the natural logit function. So it can be used to convert the logarithm of odds into a probability; the conversion from the log-likelihood ratio of two alternatives also takes the form of a logistic curve.

The logistic or sigmoid function is related to the hyperbolic tangent, by

$$2P(t) = 1 + \tanh\left(\frac{t}{2}\right).$$

Conclusions for Chaotic Systems and Financial Markets. These few examples show how even the most simple non-linear dynamics can lead to surprising and complex behaviour. Financial markets are at least complex non-linear feedback systems, but they might as well be much more complex than that. The dynamics are unknown and change over time.

This implies that financial markets will show at least the essential behaviour of non-linear, complex systems.

- A first remark is that even if we would have exact knowledge of the dynamics and precise knowledge of all relevant parameters, then still the number of stable states could be very high; and not necessarily

---

\textsuperscript{13}The logit function is the inverse of the “sigmoid”, or “logistic” function and is used in mathematics and statistics. The logit of a number $p$ between 0 and 1 is given by the formula: $\text{logit}(p) = \log\left(\frac{p}{1-p}\right) = \log(p) - \log(1 - p)$. The base of the logarithm function used is of little importance, as long as it is greater than 1, but the natural logarithm with base $e$ is the one most often used.

If $p$ is a probability then $p/(1 - p)$ is the corresponding odds in favour of the event, and the logit of the probability is the logarithm of the odds; similarly the difference between the logits of two probabilities is the logarithm of the odds ratio (OR), thus providing a shorthand for writing the correct combination of odds-ratios only by adding and subtracting: $\log(R) = \log\left(\frac{p_1}{p_2}\right) = \log\left(\frac{p_1}{1-p_1}\right) - \log\left(\frac{p_2}{1-p_2}\right) = \text{logit}(p_1) - \text{logit}(p_2)$. 

---
Figure 1: In the six graphs we see the behaviour of the solution of the logistic equation for different growth parameters $a$. Each graph has on the x-axis the time $t$, and on the y-axis $P(t)$; each graph has one value for $a$ from -1 to 1.5 in steps of 0.5; and each graph has multiple lines for different starting values $P_0$. For small positive growth parameters $a$, the logistic equation has one equilibrium state. The evolution to that state is characterized by the specific curves presented in this figure. The continuous form of the logistic equation and can be analytically integrated, and these figures represent the solutions for different initial values and different growth parameters $a$. 
May and Feigenbaum developed a specific diagram to study the equilibrium states in function of the growth parameter. This diagram is known as a “bifurcation diagram”. It shows the different equilibrium states for a given growth parameter $a$. For $a \in [-1, 1]$ the equilibrium state is 0, for $a \in [1, 3]$ there is exactly one stable state that is bigger than zero. Above 3 there are multiple stable states. The system will then oscillate between those stable states after each iteration. Above 3.56 the system is in a chaotic state. Please note that chaos does not mean random.

For $a \notin [-2, 4]$ the system is unstable and diverges.

Generally only the range where $a > 0$ is plotted, however the system is also stable for some negative values. Since the value $P_t$ becomes negative, we have to reject these values of $a$ in a model for share prices.
Figure 3: A zoom of the bifurcation diagram presented in Figure 2 on page 16. It is interesting to notice is that even in the chaotic state, there are some values that are more probable than others. Even for some parameters (e.g. $a$ around 3.84) a certain order reappears. Also interesting to notice is that these cutoff-values (the values of $a$ in which bifurcation points occur) do not come from any higher theory (yet), they are (in general) irrational numbers that can not be reduced to the well known numbers such as $\pi$ and $e$. 
Figure 4: Another characteristic of a chaotic system is that in practice it is impossible to predict the long term behaviour of the system, because the initial parameters would have to be known with an infinite precision in order to make a long term forecast. Here we have simulated the evolution of $P_t$ for $a = 3.96$ with two initial values that only differ in the seventh digit after the decimal. Already after 15 iterations the difference becomes visible in the graph, and after 23 iterations the system behaves completely different. It is also interesting to see how the system sometimes on short term seemingly displays patterns.

The reader will notice the connection with technical analysis, a common practice in portfolio management, that attributes its claim to justification in behavioural finance. Since we have evidence that patterns exist, one cannot be sure that technical analysis is impossible. However, the above graph will be an argument to use it very carefully: appearances can be deceiving!
consistent with our intuition. The model illustrated on Figure 2 on page 16 and Figure 3 on page 17 is much more simple than any economy, and still shows almost no pattern in the chaotic state.

- However, chaotic behaviour (mathematical chaos) is not the same as random behaviour. A chaotic system is in essence a deterministic system, not a stochastic system.

- A complex, non-linear system can show bifurcations: these are very abrupt shifts to totally different states. This is what happens when a bubble bursts or a Global Meltdown rocks the financial markets.

- Sometimes those systems can stay for a while in a similar mode, and then suddenly shift gear. See for example Figure 4 on page 18, between cycle 10 and 18 it seems that an increasing trend with decreasing volatility sets in, only to be suddenly broken by a deep crash in cycle 19 (also in cycles 0 to 3, 4 to 8). Think of each cyle as the annual result of a financial market.

- A last and probably very important remark is that this implies that we might have some idea about the dynamics, but no matter how good our knowledge of those dynamics is, they will never be sufficent to make predictions on the longer term. This is because even if we know the initial parameter with great precision, this will not be enough, as is illustrated in Figure 4 on page 18.

The simple models that we used to describe the essential characteristics of non-linear, complex systems are still a far cry from the complexity of a real economy. However, they already lead to very interesting results and constitute a very compelling case to let go some hopes such as finding a way to make long term forecasts. The careful reader will also have noticed that three very different examples share the same mathematics. This is done to demonstrate how generic those models are and how wide the applications are can be. Much systems of our everyday world are governed by non-linear dynamics.

2.2.5 Maslow’s theory of Hierarchy of Needs (1943) and the framing heuristic

The content of the theory is probably commonly known, however because its importance in this document we repeat here the key ideas.

To certain extend in contrast with prevailing tradition in psychology Maslow studied (what he called) “exemplary people” such as Albert Einstein, Jane Addams, Eleanor Roosevelt, and Frederick Douglass rather than mentally ill or neurotic people. He wrote that
“the study of crippled, stunted, immature, and unhealthy specimens can yield only a cripple psychology and a cripple philosophy.”
– (Maslow 1954) - p. 236.

Key Ideas. Maslow’s hierarchy of needs positions a well-defined order of importance to different needs. It is often depicted as a pyramid consisting of five levels: the lowest level is associated with physiological needs, while the highest level is associated with self-actualization needs, particularly those related to identity and purpose. Deficiency needs must be met first. Once these are met, seeking to satisfy growth needs drives personal growth. The higher needs in this hierarchy only come into focus when the lower needs are met. Once an individual has moved upwards to the next level, needs in the lower level will no longer be prioritized. If a lower set of needs is no longer being met, the individual will temporarily re-prioritize those needs by focusing attention on the unfulfilled needs, but will not permanently regress to the lower level.

Deficiency needs. The lower four layers are what Maslow called “deficiency needs” or “D-needs”: physiological, safety and security, love and belonging, and esteem. With the exception of the lowest (physiological) needs, if these ”deficiency needs” are not met, the body gives no physical indication but the individual feels anxious and tense.

Physiological needs. Physiological needs are are those needs that are the most crucial to survival on short term. Those needs are generally related to a proper functioning of the body. If these requirements are not met (with the exception of clothing, shelter (if they don’t endanger homeostasis) and sex), the human body simply cannot continue to function.

Physiological needs include:

- Breathing
- Homoeostasis
- Water
- Sleep
- Food
- Sex
- Clothing
- Shelter
Safety needs. With one’s physiological needs (to a sufficient extend) satisfied, the individual’s safety needs take over and will dominate behaviour. These needs have to do with people’s yearning for a predictable, orderly world in which injustice and inconsistency are under control, the familiar frequent and the unfamiliar rare.

Among others, these safety needs manifest themselves in such things as a preference for job security, grievance procedures for protecting the individual from unilateral authority, savings accounts, insurance policies, and the like.

For the most part, physiological and safety needs are reasonably well satisfied in the “First World.” The obvious exceptions, of course, are people outside the mainstream: the poor and the disadvantaged. They still struggle to satisfy the basic physiological and safety needs. They are primarily concerned with survival: obtaining adequate food, clothing, shelter, and seeking justice from the dominant societal groups.

Safety and Security needs include:

- Personal security
- Financial security
- Health and well-being
- Safety net against accidents/illness and the adverse impacts

Social Needs or the Love/Belonging Needs. After the fulfilment of the physiological and safety needs, the third layer of human needs is social. These needs relate to:

- Friendship
- Intimacy
- Having a supportive and communicative family

Humans need to feel a sense of belonging and acceptance, whether it comes from a large social group, such as clubs, office culture, religious groups, professional organizations, sports teams, gangs (“Safety in numbers”), or small social connections (family members, intimate partners, mentors, close colleagues, confidants). They need to love and be loved (sexually and non-sexually) by others. In the absence of these elements, many people become susceptible to loneliness, social anxiety, and clinical depression. This need for belonging can often overcome the physiological and security needs, depending on the strength of the peer pressure. Somebody who suffers from anorexia nervosa, obviously, may ignore the need to eat and the security of health for a illusion of control and belonging.
Esteem. All humans have satisfied the previous need levels, develop an important need to be respected, to have self-esteem and self-respect. Esteem needs represents the normal human desire to be accepted and valued by others. People need to engage themselves to gain recognition and have an activity or activities that give the person a sense of contribution, to feel accepted and self-valued, for example in a profession or hobby. Imbalances at this level can result in low self-esteem or an inferiority complex. People with low self-esteem need respect from others. They may seek fame or glory, which again depends on others. It may be noted, however, that many people with low self-esteem will not be able to improve their view of themselves simply by receiving fame, respect, and glory externally, but must first accept themselves internally. Psychological imbalances such as depression can also prevent one from obtaining self-esteem on both levels.

Most people have a need for a stable self-respect and self-esteem. Maslow noted two versions of esteem needs, a lower one and a higher one. The lower one is the need for the respect of others, the need for status, recognition, fame, prestige, and attention. The higher one is the need for self-esteem, strength, competence, mastery, self-confidence, independence and freedom. The last one is higher because it rest more on inner competence won through experience. Deprivation of these needs can leads to an inferiority complex, weakness and helplessness.

Maslow stresses the dangers associated with self-esteem based on fame and outer recognition instead of inner competence. Healthy self-respect is based on earned respect.

Self-Actualisation. The motivation to realize one’s own maximum potential and possibilities is considered to be the master motive or the only real motive, all other motives being its various forms. In this need level we find needs such as the search for justice, truth and understanding the world around us, problem solving, as well as the desire to express oneself in an artistic way, to create, etc. This need level will start to create a certain level of harmony between the human being’s physical and psychological existance and the surrounding world. This in the end is the gate to the next and last need level.

Self-transcendence. Near the end of his life Maslow revealed that there was a level on the hierarchy that was above self-actualization: self-transcendence [(Maslow 1971) - part VII. Transcendence and the psychology of being (pp. 259–286)].

“"The other type (transcenders?) may be said to be much more often aware of the realm of Being (B-realm and B-cognition), to be living at the level of Being, i.e., of ends, of intrinsic values (85);
Table 1: A summary of the need levels described by Maslow in (Maslow 1943). The basic needs are more towards the bottom.

<table>
<thead>
<tr>
<th>Need Level</th>
<th>Content of the Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-actualisation</td>
<td>creativity, problem solving, spontaneity, morality, lack of prejudice, acceptance of facts</td>
</tr>
<tr>
<td>Esteem</td>
<td>self-esteem, confidence, achievement, respect of others, respect by others</td>
</tr>
<tr>
<td>Love/Belonging</td>
<td>friendship, family, sexual intimacy</td>
</tr>
<tr>
<td>Safety</td>
<td>security of body, employment, resources, morality, family, health and property</td>
</tr>
<tr>
<td>Physiological</td>
<td>breathing, food, water, sex, sleep, homoeostasis, excretion</td>
</tr>
</tbody>
</table>

Maslow later did a study on 12 people he believed possessed the qualities of Self-transcendence. Many of the qualities were guilt for the misfortune of someone close, creativity, humility, intelligence, and divergent thinking. They were mainly loners, had deep relationships, and were very normal on the outside. Maslow estimated that only 2% of the population will ever achieve this level of the hierarchy in their lifetime, and that it was absolutely impossible for a child to possess these traits.

He stated also that the achievements and success of his offspring were more satisfying than the personal fulfilment and growth characterized in self-actualization.

The discussion about the transcendence need level is most interesting, however it is clear that a person who gets satisfaction from this need level does not get satisfaction from money itself. Money can play a role -as in the example above-, but in that case it is already covered by lower need levels. This need level is typically independent of money and the physical existence. Therefore, the self-transcendence level will not play a role in the rest of this document where we will focus on financial assets and portfolio

---

14Plateau Experience refers to serene and contemplative B-cognitions as opposed to climactic ones.
criticisms. While Maslow’s theory was regarded as an improvement over previous theories of personality and motivation, it had its detractors. For example, in their extensive review of research which is dependent on Maslow’s theory, Wahba and Bridgewell (Wahba and Bridgewell 1976) found little evidence for the ranking of needs Maslow described, or even for the existence of a definite hierarchy at all. Chilean economist and philosopher Manfred Max-Neef has also argued fundamental human needs are non-hierarchical, and are ontologically universal and invariant in nature - part of the condition of being human; poverty, he argues, is the result of any one of these needs being frustrated, denied or unfulfilled.

What all these studies that somehow question Maslow’s theory, is that they do not undermind the fact that people have different needs that have to be addressed at different moments. They all recognize that one need has to be safely covered in order to move on to another need. This is key for the essence of the Maslowian Portfolio Theory to be valid. The exact order of the needs is less relevant for the rest of this work. Actually, we will see further that in Target Oriented Investment Advice the order of needs is not important at all.

Further notes and link with framing. Maslow’s theory seems to be rather independent of the thinking about financial markets that was going on. A deeper analyses however indicates that this is the first description of the “framing effect” (see (Tversky and Kahneman 1981) and (Tversky and Kahneman 1986))!

Indeed 38 years before Tversky and Kahneman described the heuristic that they called “framing” (Tversky and Kahneman 1981) Maslow published his “Theory of Human Motivation” (Maslow 1943) and doing so confirmed that the essence of human motivation and behaviour is based on framing! Framing is indeed a heuristic which is inherent to human behaviour and it is a proven survival tactic.

Maslow, understood very well how the human mind works; people do indeed seldom focus on all their needs simultaneously, they will focus on that what they need the most at that moment, and then move on. This stepwise approach is essential for our survival. This heuristic is probably much older than mankind. Basically it can be observed in most (higher) animals. Most certainly the basic idea of focusing on the most urgent need only is something that stems from the Pre-Cambrian aeon.

Even for the very first animals it was important to focus on the right needs on the right time and place. A generation of a certain species that mixes focus on eating, being eaten and procreation would have been the last
of their kind.\textsuperscript{15}

This heuristic is the very basis of animal behaviour. Just as the laws of thermodynamics these are the only laws that for sure must be observed in other universes (regardless the nature of elementary particles and fundamental forces).

2.2.6 The Axioms of Von Neuman and Morgenstern

Von Neuman and Morgenstern (von Neumann and Morgenstern 1944) derived a set of necessary axioms that are needed for the utility function to exist and represent the preference structure.

Let $\Omega$ be the set of possible outcomes of a lottery, game or investment, and define the set of all possible outcomes $\Omega$ as follows.

$$\Omega = \{\omega_1, \omega_2, \omega_3, \ldots, \omega_N, \ldots\}$$

Then we can define a binary relation over $\Omega$: $\succeq$. So that we can write $\omega_k \succeq \omega_l$, meaning "outcome $\omega_k$ is preferred to or equivalent to outcome $\omega_l$".

Axiom-1 $\succeq$ is complete:
$\omega_k \succeq \omega_l$ or $\omega_l \succeq \omega_k : \forall \{k, l\}$,

in other words all alternatives are comparable (one prefers $\omega_k$ to $\omega_l$, $\omega_l$ to $\omega_k$ or is indifferent

Axiom-2 $\succeq$ is transitive:
if $\omega_k \succeq \omega_l$ and $\omega_l \succeq \omega_m$ $\Rightarrow \omega_k \succeq \omega_m \forall \{k, l, m\}$,

in other words indifference and preference are transitive

Axiom-3 Archimedean Axiom:
if $\omega_k \succeq \omega_l \succeq \omega_m \Rightarrow \exists (a, b) \in (0, 1[\right)0, 1[)$ such that $a\omega_k + (1 - a)\omega_m \succeq \omega_l$ and $\omega_l \succeq b\omega_k + (1 - b)\omega_m$.

The Archimedean Axiom works like a continuity axiom on preferences. It states that given any three lotteries strictly preferred to each other, we can combine the most and least preferred lottery ($\omega_k$ and $\omega_m$) via an $a \in ]0, 1[$ such that the compound of $\omega_k$ and $\omega_m$ is strictly preferred to the middling lottery $\omega_l$ and

\textsuperscript{15}This idea looks so simple and basic that it seems to the author that it applies to all all Eukaryotes (because they all show periods of focus on certain activities). To some extend also bacteria and even smaller organisms have the tendency to focus on different needs at different times. This means not only that this rule is probably as old as Proterozoic acon, but also that it is fundamental to life, much deeper rooted than even the brain. Whatever starting point or level of generality one will accept, it seems that this rule -that one has to focus on different needs at different moments- seems at least to be very deeply rooted, very universal and and strong argument in favour our our general approach in Maslowian Portfolio Theory.
we can combine $\omega_k$ and $\omega_m$ via a $b \in [0, 1]$ so that the middling lottery $\omega_l$ is strictly preferred to the compound of $\omega_k$ and $\omega_m$. Notice that one needs $\Omega$ to be a linear, convex structure to have the Archimedes axiom.

**Axiom-4 Independence Axiom:**

$$\forall \omega_k, \omega_l, \omega_m \in \Omega \text{ and any } a \in [0, 1]:$$

$$\omega_k \succeq \omega_l \iff a\omega_k + (1 - a)\omega_m \succeq a\omega_l + (1 - a)\omega_m.$$  

The Independence Axiom, is a little bit more troublesome. It claims that the preference between $\omega_k$ and $\omega_l$ is unaffected if they are both combined in the same way with a third lottery $\omega_l$. One can envisage this as a choice between a pair of two-stage lotteries. In this case, $a\omega_k + (1 - a)\omega_m$ is a two stage lottery which yields either lottery $\omega_k$ with probability $a$ and lottery $\omega_m$ with probability $(1 - a)$ in the first stage. Using the same interpretation for $a\omega_l + (1 - a)\omega_m$, then since both mixtures lead to $\omega_m$ with the same probability $(1 - a)$ in the first stage and since one is equally well-off if this case occurs, then preferences between the two-stage lotteries ought to depend entirely on one’s preferences between the alternative lotteries in the second-stage, $\omega_k$ and $\omega_l$.

We should note, that these axioms, as stated, are derived from N.E. Jensen (1967) and are not exactly the original von Neumann-Morgenstern (1944) axioms (in particular, they did not have an explicit independence axiom). There are, of course, alternative sets of axioms which we can use for the main theorem. One famous axiomatization was provided by I.N. Herstein and J. Milnor (1953) which is a bit more general. See (Fishburn 1982) or (Fishburn 1988) for more details.

However, using the axioms mentioned above, one can prove the following theorem:

**Theorem 1 (von Neumann Morgenstern)** Let $\Omega$ be a convex subset of a linear space. Let $\succeq$ be a binary relation on $\Omega$. Then $\succeq$ satisfies (Axioma 1), (Axioma 2), (Axioma 3) and (Axioma 4) if and only if there is a real-valued function $U : \Omega \to \mathbb{R}$ such that:

1. $U$ represents $\succeq$ (i.e. $\forall \omega_k, \omega_l \in \Omega : \omega_k \succeq \omega_l \iff U(\omega_k) \geq U(\omega_l)$)

2. $U$ is affine$^{16}$ (i.e. $\omega_k, \omega_l \in \Omega : U(a\omega_k + (1 - a)\omega_l) = aU(\omega_k) + (1 - a)U(\omega_l) \forall a \in [0, 1]$)

---

$^{16}$In general an affine combination of vectors $x_i$: $\sum_{i=1}^{N} a_i x_i$, so that $\sum_{i=1}^{N} a_i = 1$. For example the total value of a portfolio is an affine combination of the sub-portfolios, or the expected return is an affine combination of the weighted expected returns of the individual assets.
Moreover, if \( \tilde{U} : \Omega \mapsto \mathbb{R} \) also represents preferences, then there are \( a, b \in \mathbb{R} \) (where \( a > 0 \)) such that \( \tilde{U} = aU + b \), i.e. \( U \) is unique up to a positive linear transformation.

### 2.2.7 The Friedman-Savage Puzzle (1948)

The Friedman-Savage Puzzle is consists of the constatation that people buy both insurances and lottery tickets (Friedman and Savage 1948). This makes people both risk averse and risk seeking, and was in contradiction with the axiom that “marginal utility is decreasing”.

Friedman and Savage proposed the following joint hypothesis in order to solve this puzzle.

1. A person acts as if he (a) ascribed a certain level of utility to each level of wealth and (b) acts when confronted with known odds so as to maximize expected utility.

2. The utility function is as presented in Figure 5 on page 28.

These hypotheses indeed allow for both risk seeking (e.g. the purchase of lottery tickets) and risk averse (e.g. the purchase of investor grade bonds) behaviour. Some problems however remain, as was later noted by H. Markowitz (see Chapter 2.3.2 on page 34).

It might seem that this puzzle is solved by accepting that a utility function has domains where it is convex and other domains where it is concave. However this observation cannot be sufficient, because the utility function is a function in one variable (the wealth). So, one person will (depending from his wealth at the moment of the observation) be or risk seeking or risk averse. In other words, the problem is that investments and lottery tickets fall in the same domain of the utility function, and obviously the utility function cannot be convex and concave in the same place.

This puzzle can only be solved by accepting that people do have different mental accounts (separate portfolios) for different purposes and that each different portfolio can have a different risk profile.\(^{17}\)

\(^{17}\)An important corollary of this statement is that the “risk profile of an investor” does not exists. One investor has multiple risk profiles (not one), each sub-portfolio can have a different risk profile; even to such extend that in some parts of the portfolio one is risk seeking (buying lottery tickets) and in other sub-portfolios risk averse (investing in a savings account for example). Basing investment advice on a questionnaire that determines “the risk profile of the investor” is hence a dangerous mistake that will inevitably lead to wrong advice and disappointed investors.
Figure 5: The utility function as proposed by Friedman and Savage. A person with wealth level A, would buy lottery tickets that if that could bring him to level B, on the other hand, he would be inclined to insure losses that would make him drop below wealth level A.
Figure 6: Two parameter criteria try to find a ranking of possible portfolios. Other portfolios compared to the one marked by the big dot are better in quadrant D because they have lower risk and higher return, in quadrant B we have worse portfolios (with lower return and higher risk), but those in A and C cannot really be ranked at this level.

2.3 Rational Portfolio Theories (1950s and 1960s)

2.3.1 Modern Portfolio Theory (1952)

Markowitz introduced the “Modern Portfolio Theory” (Markowitz 1952a). The idea is as simple as powerful. The two most important criteria for an investment are the return and the risk. The theory of Markowitz states that preference should be given to a portfolio with higher return for similar risk, or to the portfolio that has a lower risk when returns are the same. If we would plot all possible portfolios, we would get a an upper boundary for all portfolios. This upper boundary is called the “Efficient Frontier”, all portfolios on this frontier have this in common that there is no other portfolio that offers both a higher return for a given amount of risk.

Markowitz suggested to use variance as a measure of risk. He did not really argue that this would be the best choice. His line of reasoning is first rejecting that people would/should only maximising returns. He rejects this hypothesis (p. 77) because it leads to non-diversified portfolios.

Therefore he writes “We next consider the rule that the investor does (or should) consider expected return a desirable thing and variance of return an undesirable thing.” and moves on proving that the “(E-V) rule” criterium (maximizing of Expected Return and minimizing Variance”). The theory is meant as a normative theory, he writes:

“There is a rule which implies both that the investor should
Figure 7: The mean-variance criteria visualized for the case of 2 portfolios. Two assets are assumed with respectively return 0.15 and 0.07; and standard deviation 0.2 and 0.1, with a correlation of 0.3. Please note that contrary to actual practice, Markowitz plotted expected variance on the y-axis and expected return on the x-axis.
diversify and that he should maximize expected return. The rule states that the investor does (or should) diversify his funds among all those securities which give maximum expected return."
– (Markowitz 1952a) - page 79.

Further on page 80 he argues that variance “is a commonly used measure for dispersion” and that even if one uses another measure for risk than “V” such as the standard deviation or the coefficient of variation ($\frac{\sigma}{\mu}$) that then one would find the same optimal portfolios.

The essence for Markowitz is that a portfolio selection method that would not lead to diversified portfolios (in the sense of total holdings) is unacceptable. Without any further claim of having found the best solution he proposes a solution that satisfies the condition of diversification and argues that it is wise to apply his rule because it is better than “speculative behavior” (p. 87), because it “implies the right kind of diversification for the right reason.”.

He mentions however a few conditions for his theory to be applicable.

“Two conditions –at least– must be satisfied before it would be practical to use efficient surfaces in the manner described above. First, the investor must desire to act according to the E-V maxim. Second, we must be able to arrive at reasonable $\mu_i$ and $\sigma_{ij}$”
– (Markowitz 1952a) - page 83.

**A Mathematical Formulation of the MV-criterion.** We will from now on use the more popular notation “MV-criterion” in stead of Markowitz own suggestion “(E-V) rule”.

The standard deviation of a real-valued random variable $X$ is defined as:

$$\sigma = \sqrt{AR}$$

$$= \sqrt{E[(X - E[X])^2]}$$  (definition of variance)

$$= \sqrt{E[X^2] - (E[X])^2}$$  (14)

$$= \sqrt{\int_{\mathbb{R}} (x - \mu) f(x) \, dx}$$  (for continuous distributions)

$$= \sqrt{\frac{1}{N} \sum_{k=1}^{N} (x_k - \bar{x})^2}$$  (for discrete distributions)

with:

$$\mu = \int_{\mathbb{R}} x f(x) \, dx$$

$$\bar{x} = \frac{1}{N} \sum_{k=1}^{N} x_k$$
Figure 8: This figure shows possible portfolios that consist of three assets. Portfolios are plotted with a 1% step in compositions each. Once there are more than 2 assets, the number of possible portfolios explode and generally cover a surface in the $(\sigma, R)$-plane. One also sees that the portfolios with the lowest variance include all assets. Or in other words, adding any asset that is not 100% correlated allows us to reduce volatility for a fixed return.

A simple mathematical formulation of the mean variance optimization could be the following$^{18}$.

Suppose that we want to construct a portfolio comprising of $N$ possible risky assets. A portfolio can be referred to via the weights over the different assets:

$$w = (w_1, w_2, \ldots, w_N)'$$

with condition

$$\sum_{i=1}^{N} w_i = 1 \quad (15)$$

The returns of the possible assets $R = (R_1, R_2, \ldots, R_N)'$, have expected returns $\mu = (\mu_1, \mu_2, \ldots, \mu_N)'$ and have an expected covariance matrix de-

$^{18}$Please note that this part does not follow Markowitz’ formulation, that was rather geometrical and was limited to 3 possible assets. This formulation is more general, but uses exactly the same principles.
fined by:
\[
\Sigma = \begin{pmatrix}
\sigma_{11} & \cdots & \sigma_{1N} \\
\vdots & \ddots & \vdots \\
\sigma_{N1} & \cdots & \sigma_{NN}
\end{pmatrix}
\]

Where \( \sigma_{ij} \) is the covariance between asset \( i \) and asset \( j \).

Using this formulation, it is trivial that the return, expected return and expected variance of a portfolio \( p \) are respectively defined by:
\[
R_p = w' R \\
\mu_p = w' \mu \\
\sigma_p^2 = w' \Sigma w
\]

The mean variance criterion as developed by Markowitz is now reduced to
\[
\begin{cases}
\min_w \{ w' \Sigma w \} \\
\max_w \{ w' \mu \}
\end{cases}
\]  \tag{16}

with constraint:
\[
w' \cdot 1 = 1 \tag{17}
\]

However 16 is a problem that leads to an infinite set of solutions that cannot be ordered by these two criteria: all the portfolios on the the “efficient border”.

We will therefore choose one:
\[
\mu_0 = w' \mu
\]

Hence we replace 16 with constraint 17 by:
\[
\min_w \{ w' \Sigma w \} \tag{18}
\]

with constraints:
\[
\begin{cases}
\mu_0 = w' \mu \\
w' \cdot 1 = 1
\end{cases} \tag{19}
\]

This formulation of the mean variance criterion is generally referred to as “the risk minimization formulation”. It is a quadratic optimization problem with equality restraints and is hence solved by using e.g. the method of Lagrange Multipliers.

It’s solution is given by:
\[
w = \frac{1}{a c - b^2} \Sigma^{-1} (c 1 - b \mu + (a \mu - b) \mu_0)
\]

\[\text{Please note that } \sigma_{ii} = \sigma_i^2, \text{ and that in general } \sigma_{ij} = \rho_{ij} \sigma_i \sigma_j.\]

33
where the scalars $a$, $b$ and $c$ are defined by:

\[
\begin{align*}
    a &= \mathbf{i}' \Sigma^{-1} \mathbf{i} \\
    b &= \mathbf{i}' \Sigma^{-1} \mathbf{\mu} \\
    c &= \mathbf{\mu}' \Sigma^{-1} \mathbf{\mu}
\end{align*}
\]

Of course one can also choose alternative formulations of the mean variance hypothesis, such as the “expected return maximization formulation” or even the “risk aversion formulation”.

A most interesting and theoretically very satisfying corollary of this theory is that volatility aversion of the investor can be expressed by a simple utility function where one parameter $d$ characterizes the variance aversion.

\[
U_{MPT} = \mu - \frac{\sigma^2}{d} \tag{20}
\]

### 2.3.2 Markowitz’ Customary Wealth Theory (1952)

Markowitz published in 1952 besides the article in which he introduced his Mean Variance Theory also another article entitled “The Utility of Wealth” (Markowitz 1952b). In this publication, Markowitz explained the paradox of Friedman and Savage (Friedman and Savage 1948) by noting that people aspire to move up from their current social class or “customary wealth”\(^{20}\). So, people might accept lottery-like odds in the hope of winning amounts that significantly exceed their customary wealth, especially if the amounts that can be lost are small (relative to the customary wealth level).

Considering the utility function as proposed by Friedman and Savage in Figure 9 on page 35, Markowitz remarks the following discrepancies Friedman Savage hypothesis and “common observation”, or

“behavior which not only is not observed but would generally be considered peculiar if it were. At other points on the curve the hypothesis implies less peculiar but still questionable behavior. At only one region of the curve the F-S hypothesis implies behavior which is commonly observed. This in itself may suggest how the analysis should be modified.”

– (Markowitz 1952b) - page 152.

In the enumeration below, the names of Wealth levels A, B, C and D are clarified in Figure 9 on page 35.

\(^{20}\)The “Customary Wealth” is the level of wealth to which one is used. Windfall losses or gains might to some extend be added to the actual wealth level in order to obtain the customary wealth. In absence of windfall gains or losses, the customary wealth equals the actual wealth levels to which one got used.
Persons with wealth centrally between C and D should according to the Friedman Savage hypothesis be willing to take large symmetrical bets (that allow them for example to reach D or C with equal odds. Such behaviour is not observed.

A person with wealth almost equal do D (a person that is “almost rich”, would be willing -according to the Friedman Savage hypothesis- to take a small chance of a large loss (bringing him down do C), but if won would rise him to D. Furthermore he would not insure against a loss that would bring him down to wealth level C. Indeed, a moderately wealthy person that is willing to risk a large fraction of his or her wealth at actuarially unfair odds “will arise very rarely. Yet such a willingness is implied by a utility function like [Figure 9 on this page]” (Markowitz 1952b).

A person with wealth less than C or more than D will never take a fair bet according to the Friedman Savage hypothesis. It is on the contrary observed that also poor people buy lottery tickets and rich people gamble.

One does however find that the Friedman Savage hypothesis generate quite plausible results for people that are close to wealth level A. This to-
gether with a series of questions to what extent one would accept gambles or go for the sure thing, lead us to the customary wealth hypothesis. This means that the wealth level A in Figure 9 always corresponds to the present wealth.

This and some additional considerations such as

1. generally people avoid symmetrical bets, hence the utility function must decrease faster left from point A than it increases right from point A,

2. in order to avoid the St. Peterburg’s Paradox the utility function should be bounded from above and from below,

lead Markowitz to postulate the utility function similar to the one presented in Figure 10 on page 37, where without loss of generality the customary wealth level is put equal to zero.

The level of customary wealth is, according to Markowitz, generally equal to the actual wealth, but can be distorted by recent windfall gains (or losses). This observation is of great importance and its impact should not be underestimated. This implies that one will constantly update his utility function as his wealth increases or decreases: the utility function is not a constant function, it changes as one advances in life.

This last observation is the essence of the hypothesis proposed by De Brouwer and Van den Spiegel in their solution for the “Fallacy of Large Numbers Paradox”. However this similarity was not noticed by the authors at that time.

Despite Markowitz’ own observation that his theory is only a “small modification of the Friedman-Savage analysis” (Markowitz 1952b), the Customary Wealth Theory is a leap forward compared to the utility function proposed by Friedman and Savage (Friedman and Savage 1948). A fundamental and major contribution is the fact that a utility is a quite changeable concept for an individual, and strongly linked to what the person already has attained in his or her life (Please not the similarity with framing and loss aversion). The utility function is not longer a constant function for an individual.

This work of Markowitz in 1952 is the foundation for Kahneman and Tversky’s Prospect Theory (see (Kahneman and Tversky 1979)), Lopez’s

---

21Markowitz observed that people generally would prefer gambles when the lowest possible outcome is small compared to the actual wealth level, but when this lowest outcome is of a similar order or larger than the actual wealth level, then people tend to prefer the sure thing. For example one prefers a chance one out of ten to earn $1, but prefers to get $1,000,000 for sure in stead of a one out of ten probability to get $10,000,000. Similarly when losses are involved Markowitz found similar behaviour: risk seeking for small losses, risk avverting for big losses.

22See Chapter 2.3.9 on page 49, Chapter 2.6.13 on page 71 and (De Brouwer and Van den Spiegel 2001)
Figure 10: The utility function as proposed by Markowitz. The graph shows that for gains one is first risk seeking and then loss averse as the potential gains increase. For losses one observes a similar behaviour. Furthermore, the utility function is bounded below and above, but the utility function goes steeper down for losses than it goes up for gains. Please note that this visualisation is in accordance with Markowitz’ text and only to a lesser extend in agreement with the figures in his 1952 publication.
SP/A theory (see (Lopez 1987)) as well as Shefrin and Statman’s Behavioural Portfolio Theory (see (Shefrin and Statman 2000)). On top of that Markowitz gave here a first description of the heuristics in human behaviour that later would become known as loss aversion (see (Tversky and Kahneman 1991)) and framing (see (Tversky and Kahneman 1981)).

Also noteworthy are the facts that Markowitz utility function allows for loss aversion rather than “volatility aversion” and that he remembers us to use bounded utility functions.

### 2.3.3 Roy’s Safety First Portfolio Theory (1952)

Roy argues that

> “in calling in a utility function . . . , an appearance of generality is achieved at the cost of a loss of practical significance... A man who seeks advice about his actions will not be grateful for the suggestion that he maximize expected utility.”

– (Roy 1952)

Instead, Roy argues that investors strive to minimize the probability of portfolio return falling below a subjectively designated disaster level. This behavioural maxim is referred to as the Safety First principle. This principle could be summarized as follows. Let’s note the return of the portfolio $R_p$ and the minimal desired return $R_m$ (the returns that would bring the wealth down to the subsistence level $W_s$, then the Safety First Principle is equivalent to:

$$\min_p \{ P(R_p < R_m) \} \quad (21)$$

If returns are normally distributed, then Roy’s Safety First Principle can be reduced to maximizing the “Safety First Ratio” (SF-ratio henceforth).

$$\max \{ \text{SF-ratio} \} \Leftrightarrow \max \left\{ \frac{E[R_p] - R_m}{\sigma_p} \right\} \quad (22)$$

(with $E[R_p]$ the expected return of the portfolio and $\sigma_p$ its standard deviation)

One will notice that this criterion will select the same portfolios as maximizing the Sharpe Ratio when returns are independently normally distributed.

If we focus on the case where no risk free asset exists (i.e. $\sigma_p > 0 : \forall p \in P$), where $P$ is the set of all possible (acceptable)\(^{23}\) portfolios for a certain investor and/or investment problem. Also Roy focussed on the case where

\(^{23}\)Typical restrictions would be no short selling, no derivatives, only liquid equities, only equities from certain countries/sectors, only bonds in certain currencies, etc. This can be the result of the investor’s preference, legal framework, or what is reasonably possible or accessible as well as costs involved in acquiring certain assets.
no risk free asset exist, and it is a quite reasonable assumption: there are no investments that have a zero variance over time horizons that are relevant for investments (multiple years to decades).

Figure 11: This graph illustrate how optimal portfolios could be determined for the Safety First portfolio theory. One will notice that depending on the level of the required minimal return (originally called subsistence level) any of the efficient portfolios in the sense of Markowitz can be selected. From the portfolio with the minimal variance in the limit where $R_m$ tends to $-\infty$ to the portfolio that is 100% composed of the most risky asset that is selected $R_m \geq R_{max}$.

The same two assets are considered as in Figure 7 on page 30.

In order to find the iso-SF curves in the $(\mu_p, \sigma_p)$ – plane we rewrite 21 for a certain SF level $SF_0$.

$$\mu_p = SF_0 \cdot \sigma_p + R_m$$

(23)

So, one will notice that the iso-SF-curves are straight lines in the $(\mu_p, \sigma_p)$–plane that intersect the y-axis in $R_m$. Of course we remember that this is only valid in the case that portfolio returns are normally distributed so that we could reduce equation 21 to 22.

Later there have been some interesting generalizations of the SF Portfolio Theory.

- (Tesler 1955) generalized the Safety First portfolio theory by introducing a desired probability level ($\alpha$) connected to the violation of the
minimal return. So, an investor will choose a portfolio that maximizes expected wealth \( E[W] \), subject to the constraint \( P(W \leq W_s) \leq \alpha \).

- (Arzac and Bawa 1977) extend Tesler’s model by allowing the probability level \( \alpha \) to vary. In that case, the expected utility function becomes

\[
EU = E[W] - cP(W \leq W_s) \quad (24)
\]

\[
= E[W] - cD_R(W_S) \quad (25)
\]

with \( c \) a scalar, and \( D_R(W_S) \) the decumulative distribution function. Markowitz (Markowitz 1959) agreed that this was the only functional form that was consistent with the expected utility hypothesis.

- (Elton and Gruber 1996) discuss some generalizations, a.o. the extension presented from Kataoka. Kataoka proposes that it is an investors aim to maximize the subsistence level subject to the constraint that the probability that wealth falls below the subsistence level does not exceed a predetermined \( \alpha \).

### 2.3.4 The Allais Paradox (1953)

The Allais paradox is a choice problem designed by Maurice Allais to show an inconsistency of actual observed choices with the predictions of expected utility theory. The problem arises when comparing participants’ choices in two different experiments, each of which consists of a choice between two gambles, A and B (Allais 1953). The pay-offs for each gamble in each experiment are as follows:

- **Experiment 1**
  - gamble 1A: 100% chance to win €1 mln.
  - gamble 1B: 1% chance to win nothing + 89% to win €1 mln. + 10% chance to win €5 mln.

- **Experiment 2**
  - gamble 2A: 89% chance to win nothing + 11% chance to win €1 mln.
  - gamble 2B: 90% chance to win nothing + 10% to win €5 mln.

The expected value (or average earnings per gamble are):

- **Experiment 1**
  - gamble 1A: €1,000,000
  - gamble 1B: €1,390,000
• Experiment 2
  - gamble 2A: € 110,000
  - gamble 2B: € 500,000

The Allais paradox lays herein that generally people prefer 1A over 1B while in experiment 2 they prefer 2B over 2A. Indeed, if one prefers 1A to 1B and 2B to 2A, we can write that the expected utilities of the preferred is greater than the expected utilities of the second choices (using the values above and a utility function of $U(W)$, where $W$ is (additional) wealth, we can demonstrate exactly how the paradox manifests):

$$1.00 \, U(10^6) > 0.01 \, U(0) + 0.89 \, U(10^6) + 0.1 \, U(5 \times 10^6)$$
(experiment 1)
$$0.89 \, U(0) + 0.11 \, U(10^6) < 0.9 \, U(0) + 0.1 \, U(5 \times 10^6)$$
(experiment 2)

We can rewrite the equation for experiment 2 as:

$$\Leftrightarrow 0.11 \, U(10^6) < 0.01 \, U(0) + 0.1 \, U(5 \times 10^6)$$
$$\Leftrightarrow 1 \, U(10^6) - 0.89 \, U(10^6) < 0.01 \, U(0) + 0.1 \, U(5 \times 10^6)$$
$$\Leftrightarrow 1 \, U(10^6) < 0.01 \, U(0) + 0.1 \, U(5 \times 10^6) + 0.89 \, U(10^6)$$

which contradicts the first bet.

Allais presented his paradox as a counterexample to the independence axiom (also known as the “sure thing principle” of expected utility theory).\(^{24}\)

### 2.3.5 Subjective Expected Utility Theory (SEU) (1954)

A year later Leonard Savage proposed the “Subjective Utility Theory” (Savage 1954). It combines two distinct subjective concepts: a personal utility function and a personal probability analysis based on Bayesian probability theory.

Savage proved that, if one adheres to the axioms of rationality\(^{25}\), and if one believes that an uncertain event has possible outcomes $\{\omega_k\}$ each with a utility to you of $U(\omega_k)$ then your choices can be explained as arising from a function in which you believe that there is a subjective probability of each outcome is $P(\omega_k)$, and your subjective expected utility is the expected value of the utility,

$$EU = \sum_{k=1}^{N} \{U(\omega_k).P(\omega_k)\}. \tag{26}$$

\(^{24}\)Please note the similarity to the concept “framing” (Tversky and Kahneman 1981) that was not yet defined at that moment.

\(^{25}\)This refers to a set of rationality axioms that was proved to be sufficient in order to describe a preference relation by a utility function, see (von Neumann and Morgenstern 1944).
You may be able to make a decision which changes the possible outcomes to $\tilde{\omega}_l$ in which case your subjective expected utility will become

\[
SEU = \sum_{l=1}^{N} \{U(\tilde{\omega}_l).P(\tilde{\omega}_l)\}.
\]

(27)

Which decision you prefer depends on which subjective expected utility is higher. Different people may make different decisions because they may have different utility functions or different beliefs about the probabilities of different outcomes.

Savage assumed that it was possible to take convex combinations of decisions and that preferences would be preserved. So if you prefer $\omega_k$ to $\omega_l$ and $\omega_m$ to $\omega_n$ then you will prefer $a\omega_k + (1 - a)\omega_m$ to $a\omega_l + (1 - a)\omega_n$, for $0 < a < 1$.

Experiments have shown that many individuals do not behave in a manner consistent with subjective expected utility, most prominently (Allais 1953) and (Ellsberg 1961). Savage’s response was not that this showed a flaw in his method, rather that applying his method allowed individuals to improve their decision making.

### 2.3.6 Cognitive Dissonance (1956)

In 1956, Leon Festinger introduced a new concept in social psychology: “cognitive dissonance” (Festinger, Riecken, and Schachter 1956). The idea is that when two simultaneously cognitions held by one person are inconsistent that this will produce a state of cognitive dissonance. This state produces an unpleasant feeling that will challenge our mind and will lead to changing of our believes (one of the two cognitions will be considered as wrong).

### 2.3.7 The Ellsberg Paradox Popularized by Daniel Ellsberg (1961)

The Ellsberg Paradox further demonstrates how people’s choices violate the expected utility hypothesis, even in the formulation of SEU. This paradox is named after Daniel Ellsberg who has the merit to popularize this paradox, but it was described earlier by John Maynard Keynes, and is described in this text in Chapter 2.2.3 on page 9.

### 2.3.8 The Capital Asset Pricing Model (CAPM – 1961)

The “Sharpe-Linter-Mossin mean-variance equilibrium model of exchange”, better known as the Capital Asset Pricing Model (CAPM) is used to determine a theoretically appropriate required rate of return of an asset (if that asset is to be added to an already well-diversified portfolio) in function that asset’s non-diversifiable risk. The model takes into account the asset’s sensitivity to non-diversifiable risk (also known as systemic risk or market
risk), often represented by the quantity beta ($\beta$) in the financial industry, as well as the expected return of the market and the expected return of a theoretical risk-free asset.

The model was introduced independently by

1. Jack Treynor: (Treynor 1961) and (Treynor 1962),
2. William Sharpe: (Sharpe 1964), and

All authors were building on the earlier work of Harry Markowitz about diversification and his Mean Variance Theory.\textsuperscript{26}

The CAPM is a model for pricing an individual security or a portfolio. For individual securities, we make use of the security market line (SML) and its relation to expected return and systemic risk (beta) to show how the market must price individual securities in relation to their security risk class. The SML enables us to calculate the reward-to-risk ratio for any security in relation to that of the overall market. Therefore, when the expected rate of return for any security is deflated by its beta coefficient, the reward-to-risk ratio for any individual security in the market is equal to the market reward-to-risk ratio. So we can write for an arbitrary security $k$:

$$\frac{E[R_k] - R_{RF}}{\beta_k} = E[R_M] - R_{RF}$$

The market reward-to-risk ratio is effectively the market risk premium and by rearranging the above equation and solving for $E(R_k)$, we obtain the Capital Asset Pricing Model (CAPM).

$$E[R_k] = R_{RF} + \beta_k (E[R_M] - R_{RF})$$

Where:

- $E[R_k]$ is the expected return on the capital asset
- $R_{RF}$ is the risk-free rate of interest such as interest arising from government bonds
- $\beta_k$ (the beta coefficient) is the sensitivity of the asset returns to market returns, or also $\beta_k = \frac{\text{Cov}(R_k, R_M)}{\text{Var}(R_M)}$,
- $E[R_M]$ is the expected return of the market
- $E[R_M] - R_{RF}$ is sometimes known as the market premium or risk premium (the difference between the expected market rate of return and the risk-free rate of return).

\textsuperscript{26}Sharpe received the Nobel Memorial Prize in Economics (jointly with Markowitz and Merton Miller) for this contribution to the field of financial economics.
Restated, in terms of risk premium, we find that:

\[ E[R_k] - R_{RF} = \beta_k(E[R_M] - R_{RF}) \]  

which states that the individual risk premium equals the market premium times beta.

Generally, the expected market rate of return is measured by looking at the arithmetic average of the historical returns on a market portfolio (e.g. S&P 500 or DJ EurostoXX 50).

The risk free rate of return used for determining the risk premium, is usually the arithmetic average of historical risk free rates of return and not the current risk free rate of return for the given time horizon.

**Risk and Diversification**  The risk of a portfolio comprises systematic risk, also known as undiversifiable risk, and unsystematic risk which is also known as idiosyncratic risk or diversifiable risk. Unsystematic risk is the risk associated with individual assets, and systematic risk refers to the risk common to all securities (the so called market risk). Unsystematic risk can be reduced by diversifying the portfolio by including a greater number of assets (specific risks “average out”). systematic risk within one market can not be diversified away, it is inherent to the market considered.

A rational investor should not take on any diversifiable risk, as they are avoidable and only non-diversifiable risks are rewarded within this model. Therefore, the required return on an asset (i.e. the return that compensates for risk taken), must be linked to its riskiness in a portfolio context - i.e. its contribution to overall portfolio riskiness - as opposed to its “stand alone riskiness.” In the CAPM portfolio risk is represented by variance. In other words the beta of the portfolio is the defining factor in rewarding the systematic exposure taken by an investor.

**The (Markowitz) efficient frontier and the Market Portfolio**  The CAPM assumes that the volatility-return profile of a portfolio can be optimized: an optimal portfolio displays the lowest possible level of volatility for its level of return. Additionally, since each additional asset introduced into a portfolio further diversifies the portfolio, the optimal portfolio must comprise every asset, (assuming no trading costs) with each asset value-weighted to achieve the above (assuming that any asset is infinitely divisible). All such optimal portfolios, i.e., one for each level of return, comprise the efficient frontier.

Because the unsystematic risk is diversifiable, the total risk of a portfolio can be viewed as beta.

An investor might choose to invest a proportion of his or her wealth in a portfolio of risky assets with the remainder in cash - earning interest at the risk free rate (or indeed may borrow money to fund his or her purchase
of risky assets in which case there is a negative cash weighting). Here, the ratio of risky assets to risk free asset does not determine overall return - this relationship is clearly linear. It is thus possible to achieve a particular return in one of two ways:

1. By investing all of one’s wealth in a risky portfolio,

2. or by investing a proportion in a risky portfolio and the remainder in cash (either borrowed or invested).

Figure 12: In this graph 100 portfolios are plotted composed out of 2 assets (increments of 1% in composition) and the Capital Allocation Line is drawn. One can observe how CAPM implies that the optimal portfolios are on the line through \((0, R_{RF})\) and \(p_M\).

For a given level of return, however, only one of these portfolios will be optimal (in the sense of lowest risk). Since the risk free asset is, by definition, uncorrelated with any other asset, option 2 will generally have the lower variance and hence be the more efficient of the two.

This relationship also holds for portfolios along the efficient frontier: a higher return portfolio plus cash is more efficient than a lower return portfolio alone for that lower level of return. For a given risk free rate, there is only one optimal portfolio which can be combined with cash to achieve the lowest level of risk for any possible return. This is the market portfolio \((p_M)\).
Assumptions of CAPM  All Investors:

1. try to maximize utility,
2. are rational and volatility-averse,
3. are price takers, i.e., they cannot influence prices.
4. are able to lend and borrow under the risk free rate of interest with no limitations,
5. trade without transaction costs,
6. are not taxed in any way on their investments or transactions,
7. deal with securities that are all highly divisible into small units,
8. assume all information is at the same time available to all investors.

Shortcomings of CAPM  Despite its theoretical appeal and its coherence with other theories for rational behaviour, such as expected utility theory, CAPM has many weak points.

1. CAPM assumes that asset returns are (jointly) normally distributed random variables. However this is (despite the Central Limit Theorem) an oversimplification that is unable to capture exactly those events that are of interest.\textsuperscript{27}

2. The model assumes that the standard deviation (or variance) of returns is an adequate measurement of risk. Risk in financial markets is not volatility, it is the probability of losing. The risk is a downside risk and not the upside potential.\textsuperscript{28}

3. CAPM assumes that there exists something as “expected return” in that sense that investors agree about it (homogeneous expectations assumption).

\textsuperscript{27}Indeed, the Central Limit Theorem states that the sum of identically distributed random variables will tend to a Gaussian distribution, but that convergence is weak and only valid in the centre of the distribution. Financial risks are exactly those extreme events “in the tails of the distribution”. In the literature are many references to statistical distributions in finance, one of the most successful distributions in capturing observations are the Lévy distributions, and according to the author one of the best works describing this is (Bouchaud and Potters 1997).

\textsuperscript{28}Much more relevant measures of risk can be defined such as Value at Risk, downside volatility, maximal drawdown, longest period of loss, etc. Once more a good source is (Bouchaud and Potters 1997).
4. CAPM assumes that the beliefs of investors about the probabilities match the true distribution of returns. Also this is widely studied and does not hold.\footnote{See for example the overconfidence-based asset pricing model of Kent Daniel, David Hirshleifer, and Avanidhar Subrahmanyam (Daniel, Hirshleifer, and Subrahmanyam 2001).}

5. CAPM does not appear to adequately explain the variation in stock returns. Empirical studies show that low beta stocks may offer higher returns than the model would predict; for example see (Black, Jensen, and Scholes 1972) find that “high beta” stocks on average earn less than the model predicts and “low beta stocks” earn more. Some data was presented as early as a 1969 conference in Buffalo, New York. Either these findings are rational (which saves the efficient-market hypothesis but makes CAPM wrong), or it is irrational (which saves CAPM as a theoretical model, but makes the EMH wrong; this possibility makes volatility arbitrage a possible strategy for systematically beating the market).

6. CAPM assumes that given a certain expected return investors will prefer lower risk (lower variance) to higher risk and conversely given a certain level of risk will prefer higher returns to lower ones. It does not allow for investors who will accept lower returns for higher risk. Casino gamblers clearly pay for risk, and it is possible that some stock traders will pay for risk as well.

7. CAPM assumes that there are no taxes or transaction costs, although this assumption may be relaxed with more complicated versions of the model.

8. The market portfolio consists of all assets in all markets, where each asset is weighted by its market capitalization. This assumes no preference between markets and assets for individual investors (such as home bias), and that investors choose assets solely as a function of their risk-return profile. It also assumes that all assets are infinitely divisible as to the amount which may be held or transacted.

9. The market portfolio should in theory include all types of assets that are held by anyone as an investment (including works of art, real estate, human capital...) In practice, such a market portfolio is unobservable and people usually substitute a stock index as a proxy for the true market portfolio. Unfortunately, this substitution can lead to undermines the validity of the CAPM. Further, due to the inobservability of the true market portfolio the CAPM might not be empirically testable. This was presented in greater depth in a paper by Richard Roll in 1977, and is generally referred to as Roll’s critique. See (Roll 1977).
10. CAPM assumes that investment decisions are made by taking just two dates into account, so that there is no opportunity to consume and rebalance portfolios repeatedly over time.

11. CAPM assumes that all investors will consider all of their assets and optimize one portfolio. This is in sharp contradiction with portfolios that are held by investors. People tend to have separate portfolios for separate goals (retirement, education for children, special expenses, etc.)

Some Definitions

CML Capital Market Line. The line that connects \((R_{RF},0)\) and the market portfolio. All portfolios on the CML are those in which, according to CAPM an investor should invest. It’s equation is:

\[
E[R] = R_{RF} + E[\sigma] \frac{E[R_M] - R_{RF}}{E[\sigma_M]} \quad (31)
\]

CAL Capital Allocation Line. The line that connects all portfolios that can be constructed by combining a risky portfolio \((p)\) and the riskless asset. It’s equation is:

\[
E[R] = R_{RF} + E[\sigma] \frac{E[R_p] - R_{RF}}{E[\sigma_p]} \quad (32)
\]

In CAPM, this is the line along which the investor will move his portfolios, he will allocate more to cash if he wants a safer portfolio and more to the risky portfolio if he seeks a more dynamic portfolio. Doing so, he will only balance cash and this portfolio, he will not change the portfolio itself (for example more into equities and less into bonds!).³⁰ An investor, investing according to CAPM will find that his CAL is the same as the CML.

SCL Security Characteristic Line is the line that represents the relation between the market return and the return of a specific asset \(i\) at a given time interval \(t\), it’s equation is:

\[
R_{i,t} = \alpha_i + \beta_i R_{M,t} + \epsilon_{i,t} \quad (33)
\]

Where \(\alpha_i\) and \(\beta_i\) are referred to as respectively the “alpha and the beta of the security \(i\)”.

³⁰Please note that in practice advisors will generally advise more equities and less bonds if a client seeks a more dynamic investment. This is in contradiction with CAPM, and it's called the “Asset Allocation Puzzle”, see Chapter 2.6.9 on page 66.
SML Securities Market Line shows the expected return as a function of \( \beta \). The intercept is the risk free rate of return \( R_{RF} \) and the slope is \( E[R_M - R_{RF}] \). It’s equation is hence:

\[
E[R_i] = R_{RF} + \beta_i \{ E[R_M] - R_{RF} \}
\] (34)

The SML is actually a single factor model for the price of a given asset \( i \), and so it can be considered as the graph that represents the results of the CAPM.

2.3.9 The Fallacy of Large Numbers (1963)

P. A. Samuelson builds his paradox “the fallacy of large numbers”, suggesting that time horizon is not relevant for investments (Samuelson 1963).

“that a person whose utility schedule prevents him from ever taking a specific favourable bet when offered only once can never rationally take a large sequence of such fair bets, if expected utility is maximised”

– (Samuelson 1963)

Of course “one bet” can be seen as one time unit in investing, and a series of bets is then the longer time horizon. This is very important since practitioners tend to advise more risky assets on longer investment horizons.

2.3.10 The Efficient Market Hypothesis (EMH) (1964)

The efficient-market hypothesis was first expressed by Louis Bachelier, a French mathematician, in his 1900 dissertation, “Théorie de la spéculation” (Bachelier 1900a) - also published as a book (Bachelier 1900b). His work was largely unknown among economists till P.A. Samuelson brought it to the attention. In 1964 Bachelier’s dissertation along with the empirical studies mentioned above were published in an anthology edited by Paul Coonter, and a year later, E. Fama published his dissertation.

The efficient-market hypothesis (EMH) states that the relevant financial markets are “informationally efficient”, i.e. prices on traded assets already reflect all known information. The efficient-market hypothesis states that it is impossible to consistently outperform the market by using any information that the market already knows, except through luck.

Information or news in the EMH is defined as anything that may affect prices that is un-knowable in the present and thus appears randomly in the future.

The EMH was developed by Eugene Fama at the University of Chicago Booth School of Business as an academic concept of study through his Ph.D. thesis that was published in the early 1960s. Only slowly it gained adherence, and developed a strong status both among academicians and practitioners.
Despite of its logical acceptance many business models in asset management and many individuals are trying to select portfolios that outperform the markets.

Many studies (starting with (De Bondt and Thaler 1985)) showed that markets are not efficient in the sense that patterns exist. However this does not necessarily mean that these patterns can be exploited in order to beat the financial markets. Indeed practitioners seem to have much more trouble to find strategies that continue to perform. Therefore even if markets are not efficient, but if we have no way to predict future patterns, it is not excluded that the EMH is still the best approximation to be used in analysis.

**Different forms of market efficiency** The efficient-market hypothesis requires that agents have rational expectations on top of the rational behaviour of maximizing a utility function; that on average market is correct\(^{31}\) and whenever new relevant information appears, the agents get it at the same time and update their expectations appropriately and in a rational way.

Note that it is not required that the agents are completely rational (in the sense that they show no behavioural biases). Some investors may overreact and some may underreact to new information. EMH only requires that the reactions of the investors are on average rational and correct. Individual investors’ reactions should be random and follow a normal distribution so that the effect on the market prices cannot be reliably exploited (in order to generate systematic excess return above market), especially when considering transaction costs (including commissions and spreads).\(^ {32} \)

Generally three different levels (or forms) of market efficiency are considered.

- **Weak-form efficiency**
  - One cannot generate excess returns by using investment strategies based on historical share prices.
  - Technical analysis will not be able to produce systematically excess returns, however some forms of fundamental analysis may still provide excess returns.
  - Asset prices do not show “patterns”. This implies that future price movements are determined entirely by information not contained in the actual price. Hence, prices must follow a random walk.

\(^{31}\)Please note that all agents on a certain market can be wrong, EMH only requires that the market as a whole is right (as an average).

\(^{32}\)Hence the adagio: “the market is always right”, and “there is no such thing a free lunch”.

50
• Semi-strong-form efficiency
  – Share prices adjust to new public information very rapidly and in an unbiased fashion, so that no excess return can be earned by trading on new information.
  – Semi-strong efficiency implies that neither fundamental analysis nor technical analysis techniques will be able to reliably produce excess returns.
  – To test for semi-strong-form efficiency, the adjustments to previously unknown news must be of a reasonable size and must be instantaneous. To test for this, consistent upward or downward adjustments after the initial change must be looked for. If there are any such adjustments it would suggest that investors had interpreted the information in a biased fashion and hence in an inefficient manner.

• Strong-form efficiency
  – Share prices reflect all information, public and private, and no one can earn excess returns.
  – If there are legal barriers to private information becoming public, as with insider trading laws, strong-form efficiency is impossible, except in the case where the laws are universally ignored.
  – To test for strong-form efficiency, one would test investor’s outperformance: no investors should be able to systematically outperform the market.\textsuperscript{33}

2.3.11 Models for Lifetime Portfolio Selection - 1968 and 1969

There was some literature available about portfolio selection when more than one time horizon is involved. The first multi period analysis is from (Phelps 1962) and (Tobin 1965), whereas (Mossin 1968) made the first attempt to extend the existing portfolio selection methods from one period to multiple periods. In August 1969 Paul A. Samuelson and Robert C. Merton published both in The Review of Economics and Statistics a paper about lifetime portfolio selection. (Samuelson 1969) derived the equations that govern the choice between consumption, a risky asset and a non-risky asset in “Lifetime Portfolio Selection by Dynamic Stochastic Programming” for the discrete case. (Merton 1969) did the same for the continuous-time case in “Lifetime Portfolio Selection Under Uncertainty: the Continuous Time Case”.

\textsuperscript{33}Please note that a normal distribution of a large group of managers will indeed show some lucky “star performers”. In order to reject the strong-form market efficiency, one would have to find more star performers than a normal distribution would predict.
Samuelson assumes that the individual tries (or should try) to maximize \( \int_0^T e^{\rho t} U(C(t)) \, dt \), where \( C(t) \) is the consumption and \( \rho \) is the continuous time interest rate at which can be invested and that one has a choice between a riskless asset and a risky asset whose return is governed by a Wiener process\(^{34}\). Then he studies the calculus-of-variation problem resulting from this in the continous time case. and hence maximizes

\[
\max_{\{C_t, w_t\}} \left\{ E \left[ \sum_{t=0}^T (1 + \rho)^{-t} U(C_t) \right] \right\}
\]

with the condition that \( C_t = W_t - \frac{W_{t+1}}{(1+r)(1-w_t) + w_t Z_t} \). \( r \) is the interest rate at which one can invest in the save asset and \( w_t \) is the ration of the portfolio invested in the risky asset, and \( Z_t \) is the return of the risky asset in the period \( t \).

Samuelson demonstrates that in the case of isoelastic utility functions (such as \( U = \log C \))\(^{35}\) the problem can be solved and in that case the “business man risk” is a fallacy, or otherwise stated: age in itself does not influence the risk profile. Isoelastic utility functions are problematic for the Karl Menger paradox, but in all other cases \( w_{T-j} \) becomes a function of \( W_{T-j-1} \) and the equations are till now a challenge.

Merton finds that in the same problem (and with the same notations) for in the two asset case is determined by

\[
\max_{\{C, w\}} \{ \Phi(w, C; W; t) \}
\]

Where

\[
\Phi \equiv e^{-\rho t} U(C) \frac{\partial I_t}{\partial t} + \frac{\partial I_t}{\partial W} \left[ (w(t)(a - r) + r)W(t) - C(t) \right]
+ \frac{1}{2} \frac{\partial^2 I_t}{\partial W^2} \sigma^2 w^2(t)W^2(t)
\]

In this equation \( I_t \) is used as a shorthand notation for \( I(W(t), t) \), that is defined as

\[
I(W(t), t) \equiv \max_{\{C(s), w(s)\}} \left\{ E(t) \left[ \int_t^T e^{-\rho s} U(C(s)) \, ds + B(W(T), T) \right] \right\}
\]

\(^{34}\)A Wiener process (named after Norbert Wiener is another word for to describe Brownian Motion. Essential in its definition is that the system has no memory (therefore all steps are independent from the previous ones) and that each step is governed by a Gaussian distribution. The assumption of a independently distributed returns following a Multivariate Normal Distribution is an significant simplification of reality.

\(^{35}\)An iso-elastic utility function is a utility function that displays constant relative risk aversion, Merton will call this “iso-elastic marginal utility”. Generally functions of the form \( C^{\gamma} \) (with \( \gamma \neq 0 \) and \( \gamma < 1 \)) display this property, and of course \( \log C \), the limit case with \( \gamma = 0 \). For such functions we find indeed that Pratt’s measure for relative risk aversion (see (Pratt 1964)), \( \frac{\gamma U'}{C^{\gamma+1}} \), is a constant \( (1 - \gamma) \).
Where $E(t)$ is the conditional expected value operator (conditional to the knowledge of $W(t)$).

The necessary first order conditions are then a system of nonlinear partial differential equations coupled with two algebraic equations. Such equations are hard to solve, it is only in the 1970s and 1980s that the true nature of such equations was explored more in debt by physicists in their search to understand complex nonlinear systems, see our chapter about “Chaos Theory”, Chapter 2.2.4 on page 10. The book “Strategic Asset Allocation: Portfolio Choice for Long-Term Investors” (Campbell and Viceira 2002) is completely dedicated to approximated solutions of a simplified form of this equation.

Merton confirms Samuelson’s findings that for iso-elastic marginal utility the portfolio selection decision is independent of the consumption decision. In case of log-utility the consumption decision only depends on the level of wealth.

2.3.12 An Intertemporal Capital Asset Pricing Model - 1973

In 1973 Robert C. Merton (Merton 1973) uses these findings to develop an intertemporal model for the capital market. The central idea is to write an Itô equation for the capital market.

$$dX = F(X) \, dt + G(X) \, dQ$$

Where $X$ is the vector composed of prices, return and covariances, $G(.)$ a diagonal matrix and $dQ$ the Wiener process.

The model requires knowledge of many parameters about all companies and investors and 10 assumptions (including all the assumptions needed for the CAPM) are required.

An interesting result from this model is that there are only two optimal portfolios: the riskless asset and another (market) portfolio. All investors should hence hold a combination of those two.

The model is enormously complex and impossible to integrate, but it still does not include wage income and many consumption goods whose prices can change over time.

This model of Merton was another important step and many authors worked further on this model. See for example (Breeden 1979), who tried to simplify the beta and allow for stochastic consumption.

2.4 Upcoming challengers from Psychology: the 1970s and the early 1980s

2.4.1 Availability Heuristic (1973)

Tversky and Kahneman introduced the concept: “availability heuristic”. They found that our mind constantly uses the availability bias:
“The availability heuristic is a judgemental heuristic in which a person evaluates the frequency of classes or the probability of events by availability, i.e. by the ease with which relevant instances come to our mind.”
– (Tversky and Kahneman 1973)

Another way to view this is stating that when one has to make a decision, one will be inclined to use knowledge that is already available rather than examine other (more relevant) alternatives.

### 2.4.2 More Heuristics (1974)

A year later the same authors described three heuristics that are used when making decisions when outcomes are uncertain (i.e. under uncertainty) (Tversky and Neeman 1974).

- **Representativeness.** When people are asked to judge the probability that an object or event A belongs to class or process B, probabilities are evaluated by the degree of which A is representative of B (that is by the degree A resembles B)

- **Availability.** When people are asked to assess the frequency of a class or the probability of an event, they do so by the ease with which instances or occurrences can be brought to mind. In other words, when something is fresh in our mind, it appears to us that it is more probable to occur than events that are less fresh in our mind. (already published in 1973)

- **Anchoring and adjustment.** When an anchor is available, people make estimates by starting from the initial value of the anchor. The anchor may be suggested by formulation of the problem, or it may be the result of partial computations. In both cases the adjustments made are small, and people’s guesses are strongly influenced by the anchor.

### 2.4.3 Support for CAPM and related theories – 1976

In the 1970s, there were hundreds of papers published that elaborate on CAPM. Mark Rubenstein wrote for example a paper dedicated to the support of the Generalized Logarithmic Utility Function in 1976. He argued that it captures most of the dynamics while still being able to handle mathematically. – (Rubinstein 1976a)

Later during the same year he made the link from the valuation of uncertain income streams to option pricing theory – (Rubinstein 1976b).

36 An “anchor” can be almost anything that our mind uses to hold on, or to use it as a starting point to make an estimate. It might even be totally irrelevant information.
2.4.4 Prospect Theory (1979)

Expected Utility Theory is unable to explain the Allais Paradox nor the Friedman-Savage paradox. People underweight outcomes that are merely probably relatively to outcomes that are certain. Also people discard components that are shared by all gambles under consideration (see Allais Paradox). These issues are accounted for in the theory that was constructed by Kahneman and Tversky. They called their new theory “prospect theory” (Kahneman and Tversky 1979).

In prospect theory, people make decisions based on the value of gains and losses (as opposed to the utility of total wealth), and probabilities are replaced by decision weights. The value function is relative to a reference point, and is concave for gains (risk aversion) and convex for losses (risk seeking), and is generally steeper for losses than for gains (loss aversion). The decision weights are generally lower than the corresponding probabilities, expect in the range of low probabilities.

![Value Function](image)

Figure 13: Prospect Theory assumes that human decision are not made in function of or a utility function (with total wealth as variable), but decisions are made in function of a value function (that plays the same role, but uses the outcome of the gamble as variable (i.e. the incremental wealth)).

2.4.5 Framing (1981)

Tversky and Kahneman showed that decisions display predictable shifts when the same problem is differently framed. They called this phenomenon
Prospect Theory assumes that humans do not use the real probabilities in their decision making, but a distorted function of the real probabilities. “framing” (Tversky and Kahneman 1981). A good overview of this wide and deep topic can be found in (Thaler 1999).

It is also interesting to note that the idea of framing was implicitly described by A.H. Maslow in is landmark work “A theory of human motivation” - (Maslow 1943).

2.4.6 The Volatility Puzzle (1981)

In the same year, Shiller also showed that the price volatility of stocks is higher than one could reasonably expect when looking at the volatility in dividends (Shiller 1981). This is called the “volatility puzzle”.

2.4.7 Judgemental Biases Described (1982)

Kahneman, Slovic and Tversky publish their book “Judgement under uncertainty: Heuristics and biases” and explain in 35 chapters various judgemental heuristics and biases from rational behaviour (Kahneman, Slovic, and Tversky 1982).
2.5 First Signs of Acceptance for Behavioural Finance: 1985

2.5.1 The first real evidence of market non-efficiencies

Werner De Bondt and Richard Thaler provide evidence that markets are indeed not efficient (De Bondt and Thaler 1985). In their landmark article, “Does the Stock Market Overreact?”, they show that the stock market “overreacts”, and that substantial “weak form market inefficiencies” exist. Their comparison of a portfolio of “loser stocks” with a portfolio of “winner stocks” was so clear and undeniable that it sent a shock-wave through the world of economists. Their results were so unexpected and so profound that this would become a turning point in history.

Many also see this landmark article of De Bondt and Thaler as the start of “behavioural finance”, as the beginning of what is now a widely accepted paradigm.

2.5.2 Mental Accounting

In the same year Thaler developed a new model for consumer behaviour that was built on “mental accounting” (Thaler 1985). This important concept is the emanation of framing in our behaviour. People tend to distribute their assets over different mental accounts, with each their separate goal and do not put all assets in one optimized portfolio, as the rational MPT prescribes. People have different portfolios for retirement, for children, for special projects, ... and on top of that they buy lottery tickets.

This concept of mental accounting is based on framing and the Allais paradox, and would turn out to become a cornerstone of Behavioural Portfolio Theory.

2.5.3 The Equity Premium Puzzle

The equity premium puzzle was introduced by R. Mehra and E.C. Prescott (Mehra and Prescott 1985).

“Historically the average return on equity has far exceeded the average return on short-term virtually default-free debt. Over the ninety-year period 1889-1978 the average real annual yield on the Standard and Poor 500 Index was seven percent, while the average yield on short-term debt was less than one percent. The question addressed in this paper is whether this large differential in average yields can be accounted for by models that abstract from transactions costs, liquidity constraints and other frictions absent in the Arrow-Debreu set-up[37]. Our finding is that it

[37] The Arrow-Debreu model, also known as the Arrow-Debreu-McKenzie model can be used to prove that an economic equilibrium exists where the prices are such that the
cannot be, at least not for the class of economies considered. Our conclusion is that most likely some equilibrium model with a friction will be the one that successfully accounts for the large average equity premium.\(^7\) – (Mehra and Prescott 1985)

2.6 Towards Acceptance of Behavioural Finance Among Scholars (1985 - 2000)

2.6.1 The late 1980s

Choice Theory. Tversky and Kahneman (Tversky and Kahneman 1986) write that due to framing and prospect theory, the rational theory of choice cannot be a good descriptive theory of human decision making.

The 93% Question. Also in 1986, Brinson et al. show that the variation of a portfolio is for 93% determined by the strategic choice of asset classes (Brinson, Hood, and Beebower 1986) and see also (Brinson, Singer, and Beebower 1991). This is a very important discovery for portfolio selection, because it indicates what portfolio selection actually is. It is all about getting the strategic benchmark right, i.e. the average composition of the portfolio in terms of asset classes such as shares, bonds, alternative investments, cash. All the rest such as style, geographic bias, asset manager, active or passive management, etc only account for 7% of the variation.\(^38\)

EUT modified for the Allais Paradox. Yaari, however, finds that EUT can be modified to accommodate the Allais Paradox (Yaari 1987).

SP/A Theory (1987). Lopez Lola’s SP/A theory is a general framework for human decision making rather than a portfolio theory, however it can be considered as an extension of Arzac’s Security First model.

In SP/A theory, the S stands for security, the P for potential and the A for aspiration. The concept of security is closely related to the safety in Safety First theory (avoiding low levels of wealth). Aspiration is related to a goal.\(^39\) Potential is related to the general desire to achieve higher levels of wealth (a concept that is absent in SF-theory).

\(^7\)aggregate demand equals the aggregate supply. The conditions needed are: convexity, perfect competition and demand independence.

This model is the fundamental model in the General (Economic) Equilibrium Theory. The model is named after Kenneth Arrow, Gerard Debreu and Lionel W. McKenzie

\(^38\)Interesting is to note that in the USA it is common practice to pay for a financial planning (that answers the 93% question, but in Europe the main focus is mainly on the 7%.

\(^39\)This can be considered as a generalization of the SF-theory, however it has necessarily deep implications. Once we start talking about a goal, necessarily one will be confronted by the fact that humans have multiple investment goals and that the portfolio necessarily
In all Safety First models risk is represented by the probability of falling below a certain subsidence wealth level, \( W_s \). This probability is best expressed by the complementary cumulative distribution function (or a decumulative probability function) \( D \), defined as \( D(x) = P(W \geq x) \).

Assume the discrete setting, so that after a certain horizon an investment can have \( M \) different states, that occur with a probability \( p_i \) so that \( p_i = P(W_i) \; \forall i \in \{1, 2, \ldots, M\} \). Assume without loss of generality that the wealth levels are ordered so that \( W_1 \leq W_2 \leq \cdots \leq W_M \). Then one will notice that the expected wealth after the experiment equals \( E[W] = \sum_{i=1}^{M} p_i W_i = \sum_{i=1}^{M} D_i(W_i - W_{i-1}) \), with \( W_0 = 0 \). One will notice that \( D_1 = 1 \), i.e. one will receive \( W_1 \) with certainty, and can additionally get \( W_2 - W_1 \) with a probability \( D_2 \), etc.

SP/A theory now includes two aspects that influence human decisions: hope for potential and fear (or the need for security).

Fear is the human need for security and will lead to make calculations not with \( p_i \) but with modified probabilities that overweight the probabilities of bad outcomes (so assume that the probability of low gains (low \( i \)) corresponds to higher values than the actual probabilities. This is the concern for security and it can be formulated as follows.

\[
h_s(D) = D^{1+q_s}
\]

where the \( s \) refers to security and with \( q_s > 0 \).

Hope for potential will result on the contrary in overweighting positive outcomes (so the higher \( i \)). This hope for potential will result in SP/A theory in another distortion of the ccdf.

\[
h_p(D) = 1 - (1 - D^{1+q_p})
\]

where the \( p \) refers to potential and with \( q_p > 0 \). One will notice that because of the overweighting of the higher results, \( h_p(D) \) will stochastically dominate \( D \).

---

The “complementary cumulative distribution function” (ccdf) is sometimes referred to as “decumulative distribution function” (ddf) or “survival function” (mainly in survival analysis, the discipline that studies death in biological systems or failure in mechanical systems). We will write the ccdf for a stochastic variable \( X \) as \( D_X \), and it is defined by

\[
D_X(x) = P(X > x) = 1 - F_X(x)
\]

where \( F_X(x) \) is the cumulative distribution function (cdf), defined as \( F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t) \, dt \), where \( f(t) \) is of course the probability density function (pdf).

Please note that the subscript \( X \) is in this discussion omitted in the notation for the ccdf, \( D_X \), in order not to overload the notations.
Then SP/A theory combines simply this need for security and hope for potential by using the following distorted ccdf.

\[ h(D_X) = a h_s(D_X) + (1 - a) h_p(D_X) \]  
(35)

The individual person presented with a decision where outcome is uncertain will then according to SP/A theory not base his decisions on \( E[W] \), but on \( E_h[W] \), which is obtained by replacing all probabilities \( p_i \) by \( \tilde{p}_i = h(D_{i+1}) - h(D_i) \). So, here SP/A theory differs from EUT by using a rank dependent formulation (i.e. replacing \( D_X \) by \( h(D_X) \)).

Furthermore Lopez postulates that decisions when outcome is uncertain are based on two functions.

- the modified expected wealth, \( E_h[W] \)
- \( D(A) = P(W > A) \), or the probability that the outcome will exceed a certain aspiration level.

Hence the function to evaluate possible outcomes

\[ U = U (E_h[W], D(A)) \]  
(36)

is a monotonous increasing function, conform the Arzac Bawa generalization of the SF theory.

This utility function could for example be a Cobb-Douglas\(^{42}\) function.

\[ U_{SPA-CD} (E_h[W], D(A)) = (D(A))^{1-\gamma} E_h[W]^{\gamma} \]  
(37)

Now, it becomes also clear how SP/A-theory could be regarded as an extension of SF-theory (where humans base their decisions on \( E[W] \) and \( P(W \leq W_s) \)). The expected wealth is in the SP/A theory distorted by hope and fear, and the subsistence level is that theory a more general aspiration level (so, it can be a disaster level such as in SF theory, but it can also be another level). In the limit case where \( q_s = q_p = 0 \), the SP/A model is equivalent to the Arzac-Bawa model.

It is worth to mention here that “risk” is multidimensional in the SP/A model. Risk is characterized by 5 parameters.

1. \( q_s \) is a measure for the strength of fear, or the need for security;
2. \( q_p \) is a measure for the strength of hope, or the desire for potential;
3. \( A \) is the value of the aspiration level;

\(^{42}\)When Cobb and Douglas introduced their functional form in (Cobb and Douglas 1928), they presented it as a functional for of the production function of a company. The production level is characterized by \( Y = A L^\alpha K^\beta \), where \( Y \) represents the production level, \( A \) the total factory productivity, \( L \) the labour input and \( K \) the capital input. In the special case where \( \alpha + \beta = 1 \), there are “constant returns of scale”.

60
4. $a$ is a measure for the strength of fear relative to the strength of hope;

5. at least one other parameter that is a measure for the strength of the desire to reach the aspiration level relative to fear and hope in 36.

The aspiration level $A$ merits a closer look. The fact that the idea of a subsistance level of wealth (such as in Roh’s Safety First theory) is abandoned in favour of an aspiration level is not a big step in itself. It does not exclude that there is only one aspiration level, but it at least semantically a step closer to a problem setting with multiple goals.

**Further Evidence for non rational heuristics in human decision making.** De Bondt and Thaler report further evidence supporting the overreaction hypothesis (De Bondt and Thaler 1987). Samuelson and Zeckhauser find evidence of the “status quo bias” (the tendency to prefer the omission of action over action - a heuristic that reduces future regret by hindsight) (Samuelson and Zeckhauser 1988).

### 2.6.2 Evidence of Analyst overreaction

De Bondt and Thaler already had presented evidence (De Bondt and Thaler 1985) and (De Bondt and Thaler 1987) that financial markets overreact. However one could still expect that this was a bias caused by the non-professional investors. Therefore the same authors investigate the behaviour of professional analysts and find evidence that also they overreact (De Bondt and Thaler 1990). They conclude that the analysts earnings forecasts as tracked by the Institutional Brokers Estimate System (IBES) are indeed “consistent with generalized overreaction”. Further they show that the earning chances foretasted by the analysts are significantly more extreme than actual realizations. They conclude that

“the earnings forecasts are simply too extreme to be considered rational”

– (De Bondt and Thaler 1990)

### 2.6.3 The Discussion about the Efficiency of the Stock Markets

Since J.M. Keynes’s statement that most investor’s decisions “can only be taken as a result of animal spirits – of a spontaneous urge to action rather than inaction, and not as the outcome of a weighted average of benefits multiplied by quantitative probabilities” (Keynes 1936) and the formulation

43The findings of De Bondt and Thaler are somehow questioned by (Abarbanell and Bernard 1992), however also their findings and conclusions lead also to systematic deviations from rational behaviour for analysts. They however report both over and underreaction.
of the EMH many research has been devoted to the nature or the stochastics related to financial markets. The conclusions of Fama’s famous survey (Fama 1970) is that most of these studies are unable to rejected the “efficient market hypothesis”.

In 1988 Andrew Lo and Mackinlay compare variance estimators derived from data sampled at different frequencies and conclude that

“The random walk model is strongly rejected for the entire sample period (1962-1985) and for all subperiod for a variety of aggregate returns indexes and size-sorted portofolios. Although the rejections are due largely to the behavior of small stocks, they cannot be attributed completely to the effects of infrequent trading or time-varying volatilities. Moreover, the rejection of the random walk for weekly returns does not support a mean-reverting model of asset prices.”
– (Lo and MacKinley 1988)

2.6.4 Loss Aversion and endowment effect

Kahneman et al. reported evidence that people are “loss averse” and that the “endowment effect”44 persists even in settings with opportunities to learn and conclude that they are fundamental characteristics of human decision heuristics (Kahneman and Thaler 1990)

2.6.5 And more Contributions to the Behavioural Paradigm

Many other articles and books show that more and more scholars become convinced that significant deviations from rational behaviour occur. Gilovich wrote the book “How We Know What Isn’t So”, about the fallability of human decision making and learning (Gilovich 1991).

Tversky and Kahneman present a reference dependent model of riskless choice that is constructed on loss aversion (Tversky and Kahneman 1991).


Banerjee presents a simple model of herd behaviour (Banerjee 1992).

2.6.6 Cumulative Prospect Theory (1992)


44The endowment effect (also known as divestiture aversion) is a hypothesis that people value a good or service more once their property right to it has been established. In other words, people place a higher value on objects they own than objects that they do not. In one experiment, people demanded a higher price for a coffee mug that had been given to them but put a lower price on one they did not own.
It uses cumulative decision weights (and not longer the separable decision weights), it works for certain as well as for risky prospects with any number of outcomes, and it allows for different weighting functions for losses and gains.

Just as Prospect Theory this theory has no ambitions to be a normative theory. It strictly seeks to describe human decision making under uncertainty. Its assumptions and potential are quite similar to Prospect Theory, however theoretically it is an important improvement. The main modification to Prospect Theory is that, as in rank-dependent expected utility theory, cumulative probabilities are transformed, rather than the probabilities itself. This leads to the aforementioned overweighting of extreme events which occur with small probability, rather than to an overweighting of all small probability events. The modification helps to avoid a violation of first order stochastic dominance and makes the generalization to arbitrary outcome distributions easier.

This theory “predicts”:

- risk aversion for gains and risk seeking behaviour for losses of high probability, and
- risk seeking for gains and risk aversion for losses of low probability.
- The theory is also successful in explaining the equity premium puzzle, the endowment effect, status quo bias, asset allocation puzzle, etc.

Its basic assumptions are

- Utility is a function of gains and losses not of the total final wealth.
- The shape of the utility function (or value function as they call it) is convex for gains (risk averse) and concave for losses (risk seeking), with a greater sensitivity for losses than for gains (this phenomenon is also known as loss aversion, see for example (Barberis and Huang 2001)).
- A final aspect of the Cumulative Prospect Theory is the non-linear transformation of probabilities.

They also reported confirming experimental evidence of their theory.

2.6.7 A Solution for the Equity Premium Puzzle (1995)

When one studies the returns of equities over the last century, and compares them with what EUT would predict, one finds that the risk premium is far higher than the theory would predict. This paradox is referred to as the “equity premium puzzle”. Benartzi and Thaler proposed a solution for the equity premium puzzle based on behavioural concepts (Benartzi and Thaler
Figure 15: Cumulative Prospect Theory assumes (similar to Prospect Theory) that human decision are not made in function of or a utility function (with total wealth as variable), but decisions are made in function of a value function (that plays a similar role, but uses the outcome of the gamble as variable (i.e. the incremental wealth)).
Figure 16: Cumulative Prospect Theory assumes that humans do not use the real probabilities in their decision making, but a distorted function of the real probabilities. The difference with Prospect Theory lies herein that the conditions are posted on the cumulative probability distribution.

They based their reasoning on loss aversion and the tendency to monitor too frequently one’s wealth. The authors named this combination of loss aversion and monitoring on a too short time frame “myopic loss aversion”.


In the literature there are thousands of papers that refer to the milestone articles of Linter, Mossin, Markowitz, Breeden and Merton. The interest in those theories never faded and in the late 1990s there are some interesting generalizations.

Barberis considers the role of the estimation risk in a multi-period setting (Barberis 1995) but under the restrictive constraint that an investor cannot adapt his strategy as he learns more about the parameters of the distribution.

Two years later Brennan will study the influence of “learning” – published later in (Brennan 1998). He assumes that the volatility is known and

---

45Too frequently monitoring returns of one’s wealth is very common among investors. It seems to the author that it is another manifestation of framing and loss aversion. People tend not to see the big picture (wealth on long term) or at least have the illusions that they can maximize it by monitoring many short time frames.
constant, investors are able to trade continuously, and the risky asset follows diffusion process. By assuming that the expected return is not constant over time and that people are able to learn what the temporary value is he finds that in the discrete time case the uncertainty over the return increases and that this leads to a reduction in the fraction of the portfolio allocated to the risky asset.

Tong Kim and Eward Omberg (Kim and Omberg 1996) studied a simple model with only two assets (a risk free and a risky asset) that is essentially a mean reverting Wiener process and they solve it for HARA utility functions.\footnote{HARA utility functions show Hyperbolic Absolute Risk Aversion. The Absolute Risk Tolerance $T(W)$ is defined as $T(W) = \frac{\partial wU}{\partial wU^2}$. For HARA utility functions $T$ is per definition a linear function of wealth:}

They find that even such a simple model with a very straightforward utility function for only two assets leads to very complex nonmyopic behaviour. Some investors will hold more of the risky asset than the myopic investor, some less, some hedge, some leverage.

Also interesting empirical research is (Brennan, Schwartz, and Lagnado 1997) “Strategic Asset Allocation”. In this paper they find via empirical tests that the optimal composition the portfolio of an investor with a long time horizon is different than the optimal portfolio for a short time horizon. This empirical work adds arguments to our discussion of the Fallacy of Large Numbers – Chapter 2.6.13 on page 71. A second important finding is that “Out of sample simulation results provide encouraging evidence that the predictability of asset returns is sufficient for strategies that take it into account to yield significant improvements in portfolio returns.”

2.6.9 More Evidence

Momentum Strategies Pay Off. Chan and Lakonishok found that both price and earnings momentum strategies were profitable (Chan, Jegadeesh, and Lakonishok 1996). This implies that there is underreaction to new information.

The Asset Allocation Puzzle. Canner, Mankiw, and Weil publish “An Asset Allocation Puzzle” (Canner, Mankiw, and Weil 1997) and show that in general advisors tend to advice to invest a larger portion in more risky assets when one seeks more dynamic portfolios. This is in contradiction with CAPM where investors would keep all one optimal market portfolio,
but adapt the portion of cash relative to that portfolio in order to modify the riskiness of their portfolio.

Caginalp and Laurent find patterns and confirm that they are statistically significant even in out-of-sample testing, and report:

“Using two sets of data, including daily prices (open, close, high and low) of all S&P 500 stocks between 1992 and 1996, we perform a statistical test of the predictive capability of candlestick patterns. Out-of-sample tests indicate statistical significance at the level of 36 standard deviations from the null hypothesis, and indicate a profit of almost 1% during a two-day holding period. An essentially non-parametric test utilizes standard definitions of three-day candlestick patterns and removes conditions on magnitudes. The results provide evidence that traders are influenced by price behaviour”
– (Caginalp and Laurent 1998)

A Behavioural Explanation for Market Crashes. Bikhchandani et al. argue that the theory of “observational learning” (in particular “informational cascades”) can help to explain stock market crashes (Bikhchandani, Hirshleifer, and Welsch 1998).

A Model Including Over- and Underreaction. Barberis et al. present a model of investor sentiment that displays underreaction of stock prices to news such as earnings announcements and overreaction of stock prices to series of good or bad news (Barberis, Shleifer, and Vishny 1998).

2.6.10 Critics to Behavioural finance (1998)

Fama defends in his third review paper the EMH (Fama 1998). He argues that the apparent overreaction of stock prices to information is as common as underreaction and that they therefore cancel out.

However, even if that is true, over- and underreaction appear in different circumstances and hence in different time intervals. Therefore the market will display periods with overreaction and periods with underreaction: periods where the effects are not cancelled out.

2.6.11 An avalanche of evidence, theorems and books (1998 - 2000)

Further evidence was provided:

- Odean tested and found evidence of the “disposition effect” (Odean

47The “disposition effect” is the tendency to “sell winners too soon and ride losers too long”.

67
Odean brings further evidence that the trading volume in equity markets is excessive and that overconfidence would explain the phenomenon. He also found evidence of the disposition effect (Odean 1999).

- over- and underreaction (Hong and Stein 1999),
- herding (Nosfinger and Sias 1999),
- “affect heuristic” (Finucane et al. 2000),
- Lee and Swaminathan showed that past trading volume provides a link between “momentum” and “value” strategies, and argues that these findings help to reconcile the intermediate-horizon underraction and the long-horizon overreaction effects (Lee and Swaminathan 2000).

A good proof of overconfidence was presented by Barber and Odean (Barber and Odean 2001). From other psychological research we know that men are more overconfident than women (this is even more pronounced in the men dominated areas such as finance). And since we know from theoretical models that overconfidence leads to excessive trading, we could expect to find men having a higher portfolio turnover than women. Barber and Odean indeed found that men trade 45% more than women. But they found more: they found that investors decrease their portfolio performance (compared to the original composition) with each trade!

- Huberman provides evidence that people are prone to invest in the familiar (known) stocks and ignore any portfolio theory (Huberman 2001)

Different books were edited:

- “Simple Heuristics that Make Us Smart” (Gigerenzer, Todd, et al. 1999),
- “Inefficient Markets: An Introduction to Behavioral Finance” (Shleifer 2000),
- “Beyond Greed and Fear” (Shefrin 2000) (a good introduction to Behavioral finance that is now a standard work),
- “Irrational Exuberance” (Shiller 2000) (including an interesting claim that the US stock market is overvalued),

48 The “affect heuristic” concerns “goodness” and “badness”. Affective responses to a stimulus occur rapidly and automatically and strongly influence our choices.
● “Choices, Values and Frames” (Kahneman and Tversky 2000) (a selection of their work),

And some new theories and models were presented:

● Daniel et al. propose a model for security markets that is based on overconfidence of the investors (about the precision of private information) and biased self-attribution of the achieved performance, that in its turn will influence the overconfidence. (Daniel and Subrahmanyam 1998). Their model leads to significant over- and underreactions.

● Rabin constructs a theorem that shows the inadequacy of EUT (based on wealth) to explain risk aversion (Rabin 2000), “Diminishing Marginal Utility of Wealth Cannot Explain Risk Aversion”:

● in the book “Bounded Rationality: The Adaptive Toolbox”, the concept “bounded rationality” is presented as a model for how people make decisions, the toolbox is a set of heuristics (Gigerenzer and Selten 2001),

● Barberis et al. incorporate prospect theory in a model of asset prices in an economic model (Barberis, Huang, and Santos 2001)

● /* — higher moment portfolio theory — */

2.6.12 Behavioural Portfolio Theory (BPT – 2000)

“Behavioural Portfolio Theory”, BPT henceforth, as presented by Shefrin and Statman (Shefrin and Statman 2000) is a portfolio theory that breaks an important tradition: it allows for investors to have “mental accounts”:49. They present their theory as a strictly descriptive theory, and start from earlier theories: Customary Wealth Theory (see Chapter 2.3.2 on page 34, and SP/A theory (see Chapter 2.6.1 on page 58).

BPT-SA The authors present first a single account version of their theory (BPT-SA henceforth) as a special version of the SP/A theory. The efficient frontier for the BPT-SA theory is obtained by maximizing $E_h(W)$ for fixed $P(W \leq A)$ (with A, the aspiration level and all other notations similar to SP/A theory; see Chapter 2.6.1 on page 58).

49A mental account is part of the portfolio that is considered separately (for example these are the savings for my daughter), they can but not necessarily are also legally separated on different accounts at different brokers or banks. It is easy to understand why this happens: people overlook correlations and suffer from “framing” (see (Tversky and Kahneman 1981) and (Tversky and Kahneman 1986)), they do not look at all their holdings and put them in one portfolio as mean-variance optimization would suggest (see Chapter 2.3.1 on page 29).
Shefrin and Statman present then three theorems that characterize the efficient BPT-SA portfolios, the mean-variance efficient portfolios and then in a third theorem prove that generally optimal portfolios for one theory are not optimal in the other theory.

In BPT-SA, investors with a high aspiration level can find that the most risky bet (that has the highest pay-off) might be the only asset to buy. This is not because they are risk-seeking, but because this single bet provides the highest probability to reach their high aspiration level.

Another interesting remark of Shefrin and Statman is that the SP/A framework is similar to a value-at-risk framework (VaR henceforth). Indeed, in both approaches optimization is finding a trade-off between expected return or wealth and the probability of ending up below a certain threshold.

**BPT-MA** The ingeniousity of (Shefrin and Statman 2000) shows when they present the “multiple account version” of their theory: “BPT-MA”. The insight that investors have multiple goals and hence multiple sub-portfolios was an important breakthrough. They argue that investors generally combine low and high aspiration levels, i.e. investors have multiple “mental accounts”, because they overlook covariances. They argue that this is a feature of prospect theory, as developed by (Kahneman and Tversky 1979) and provide further arguments why it is reasonable to believe that people use mental accounts.

- (Tversky and Kahneman 1986) present general evidence that humans do not handle joint probability distributions at all, in stead they simplify decisions by breaking up the problem in mental accounts.

- (Yoram Kroll and Rapoport 1988) find experimental evidence that people do not take covariances into account when constructing portfolios.

- (Jorion 1994) finds that institutional investors construct portfolios as layered pyramids that therefore are not mean-variance optimal. Instead he finds that for the same variance on average 40bps more return can be found in a mean variance optimal portfolio.

Using this evidence, Shefrin and Statman formulate the multiple account version of BPT for 2 aspiration levels (a low, but important level and a high but less important level) and discrete returns. They assume a Cobb-Douglas utility function for both mental accounts.

For both levels $i$, the utility function is:

$$U_i = D(A_i)^{1-\gamma} E_h(W_i)^{\gamma}$$ (38)
where \( i \) is 1 and 2 (for aspiration level 1 and 2); and \( \gamma_i \in [0, 1] \). \(^{50}\)

One will notice that if \( \gamma = 0 \), that then Equation 38 on page 70 is equivalent with Equation 21 on page 38 of the Safety First theory. If \( \gamma = 1 \) the expected utility is close to a distorted expectation of wealth, something that probably should be rejected on basis of the St Petersburg Paradox (see Chapter 2.1.3 on page 5).

The combined utility function is such that it is zero when \( A_1 \) is not reached, but it will be higher than zero when \( A_1 \) is attained, even if \( A_2 \) is not fulfilled. This will force the first investable euro to be invested in the mental account 1.

\[
U = K_1 \left[ D(A_1)^{1-\gamma_1} E_h(W_1)^{\gamma_1} \right] \left[ 1 + K_2 D(A_2)^{1-\gamma_2} E_h(W_1)^{\gamma_2} \right]
\] (39)

where \( K_i \) is a constant that reflects the relative weight attached to each aspiration level.

The results of this approach are that investors will match mental accounts with goals, and that the different mental accounts are not integrated in the sense that covariances are considered. Indeed it can even be optimal to have a short position in a certain security in one mental account and a long position of the same security in the other portfolio!

"Portfolios within BPT-MA resemble layered pyramids where each layer (i.e., mental account) is associated with a particular aspiration level."

– (Shefrin and Statman 2000)

Once more Shefrin and Statman find that those portfolios are generally not mean-variance optimal. Indeed the most optimal securities seem to be bonds and lotteries, and one will hold both.

Please note that, anywhere else where we refer to BPT, BPT-MA is understood.

2.6.13 Fallacy of Large Numbers Revisited (2001)

De Brouwer and Van den Spiegel solve the "fallacy of large numbers puzzle"\(^{51}\) (De Brouwer and Van den Spiegel 2001). They argue that a utility function is updated once a certain level of wealth is attained, with the actual

\(^{50}\)Please note the difference between 38 and the article of Shefrin and Statman (Shefrin and Statman 2000), they use "P" in stead of \( D(A_i) \), and define P as the "probability of falling short of the aspiration level". This would make the utility an increasing function of the shortfall probability . . . so utility would decrease if the probability of reaching the aspiration level increases.

Another interesting point to make is that if in the right hand side one has the expected value of wealth, that then one should have -in our oppinion- in the left hand side also expected utility and not just utility.

\(^{51}\)See (Samuelson 1963) and chapter 2.3.9
wealth level as a reference point. They find that a simple piecewise linear utility function that shows loss aversion relative to the actual wealth level is able to produce a counterexample for Samuelson’s puzzle.

The importance of this result is that when one advises (or selects) an investment portfolio, that indeed more volatile portfolios can be selected when the investment horizon is longer.

The authors did not point to the intimate relation between their solution and existing theories like customary wealth theory, SP/A and prospect theory, but one will notice that this solution has intimate links with behavioural approaches. Actually mainly the wording used is different. De Brouwer and Van den Spiegel generalized the the concept of utility function towards a function that does not solely depend on the total wealth, but also on a certain reference level (that is updated when a new level of wealth is attained) and profits and gains relative to that level. This approach is similar to the customary wealth theory.

Richard Thaler uses Prospect Theory to come to the same conclusions (Thaler 2004) and uses almost the same piecwise linear utility function, although he calls it “value function” in line with the usual terminology in Prospect Theory.

2.7 2002: an Excellent Behavioural Finance Year

2.7.1 Nobel Price in 2002 for Behavioural Finance

Daniel Kahneman won the “2002 Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel” for his work on Prospect Theory. One can of course understand that this also honours Amos Tversky who died on June 2nd, 1996. Also noteworthy is that Daniel Kahneman got the Nobel Prize for economics while being a psychologist.

The last decades (especially since the landmark publication of De Bondt and Thaler (De Bondt and Thaler 1985)) already convinced scholars around the globe that Behavioural Finance was a paradigm with many valid aspects, but 2002 is an important breakthrough for Behavioural Finance with the larger public. Many practitioners continue to use the models of 1950a and 1960s, but they are more and more aware that J.F.M Keynes and A. Smith were right: financial markets are governed psychological phenomenon.

2.7.2 More Heuristics

In the introduction of “Heuristics and Biases: The Psychology of Intuitive Judgement (Gilovich and Griffin 2002), the authors identify

- six general purpose heuristics (affect, availability, causality, fluency, similarity and surprise),
six special purpose heuristics (attribution, substitution, outrage, prototype, recognition, choosing by liking and choosing by default),

- representativeness was replaced by attribution-substitution (this is the prototype and similarity heuristic)

- anchoring and adjustment were replaced by the affect heuristic

2.7.3 And More Evidence

Slovic et al. find evidence for the affect heuristic: (Slovic et al. 2002).

2.8 Efforts to Bring the Major Paradigms Closer to Each Other

Recently there have been some interesting papers that try to bring both paradigms closer to each other, for example (Lo 2004) and (Van den Spiegel 2005). Both papers underline the fact that there are periods where both one of the two paradigms seems to be the best adapted, but that the dynamics seem to change at unpredictable time intervals.

This is indeed what one would expect from systems that are in a deterministic chaos regime.

3 An Interpretation of the Milestones

In Table 2 on page 75, we give an overview of the key developments in a chronological way. In this table we tried to show the most important developments in favour of both rational and behavioural school. When a discovery or theory is in favour of a certain school, it appears in under this school’s heading. For example, Blaise Pascal’s “expected return” criteria was definitely a rational selection method and therefore is in the column “Rational School”.

It seems that challenging existing theories is an essential element in scientific development. Formulating paradoxes is one way of challenging existing theories and is in many cases a necessary catalyst to come to a new and better theory. Therefore the following table will focus on theories and paradoxes.

This is, according to the author, an example of a simple but important truth. Scientific progress stalls once we are deeply convinced about the ruling paradigm, once we start “believing” in stead of ”thinking”. Having an open mind for disconfirming evidence is essential for scientific development.\footnote{This remark will to the reader become more convincing when when the many cognitive biases (especially belief perseverance) are taken into account.}
Also interesting to notice is how much interwoven the development of both schools is.

<table>
<thead>
<tr>
<th>Period</th>
<th>Rational School</th>
<th>Behavioural School</th>
</tr>
</thead>
<tbody>
<tr>
<td>XVII</td>
<td>1654 (Pascal 1654): Expected Return</td>
<td>(Bernoulli 1713b): St. Petersburg Paradox</td>
</tr>
<tr>
<td>XVIII</td>
<td>1713</td>
<td>(Bernoulli 1738): concept utility function &amp; EUT</td>
</tr>
<tr>
<td></td>
<td>1738</td>
<td>(Smith 1759): behavioural economic actors</td>
</tr>
<tr>
<td></td>
<td>1759</td>
<td>(Mackay 1841): anecdotal evidence of non rational behaviour</td>
</tr>
<tr>
<td>XIX</td>
<td>1841</td>
<td>(Keynes 1921) (Ellsberg 1961): The Ellsberg Paradox</td>
</tr>
<tr>
<td>XX</td>
<td>1921</td>
<td>(von Neumann and Morgenstern 1944): axioms for the utility function</td>
</tr>
<tr>
<td></td>
<td>1940s</td>
<td>(Maslow 1943): Hierarchy of Needs</td>
</tr>
<tr>
<td></td>
<td>1940s</td>
<td>(Friedman and Savage 1948): Friedman-Savage puzzle</td>
</tr>
<tr>
<td></td>
<td>1950s</td>
<td>(Markowitz 1952a): mean variance criterion, or MPT</td>
</tr>
<tr>
<td></td>
<td>1950s</td>
<td>(Markowitz 1952b): Customary Wealth Theory</td>
</tr>
<tr>
<td></td>
<td>1950s</td>
<td>(Roy 1952): Safety First Theory</td>
</tr>
<tr>
<td></td>
<td>1950s</td>
<td>(Allais 1953): Allais Paradox</td>
</tr>
<tr>
<td></td>
<td>1960s</td>
<td>(Samuelson 1963): the fallacy of large numbers puzzle</td>
</tr>
<tr>
<td></td>
<td>1960s</td>
<td>(Fama 1964): The EMH as in (Bachelier 1900b)</td>
</tr>
<tr>
<td></td>
<td>1970s</td>
<td>(Kahneman and Tversky 1979): Prospect Theory</td>
</tr>
<tr>
<td></td>
<td>1980s</td>
<td>(Shiller 1981): volatility puzzle</td>
</tr>
</tbody>
</table>

continued on next page . . .
- continued from previous page.

<table>
<thead>
<tr>
<th>Period</th>
<th>Rational School</th>
<th>Behavioural School</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990s</td>
<td>(Yaari 1987): include the Allais Paradox in the EUT</td>
<td>(Tversky and Kahneman 1981): framing, and later (Thaler 1985): mental accounting</td>
</tr>
<tr>
<td></td>
<td>(De Bondt and Thaler 1985): first evidence of non rational behaviour in financial markets</td>
<td>(Mehra and Prescott 1985): equity premium puzzle</td>
</tr>
<tr>
<td>XXI 2000s</td>
<td>(De Brouwer and Van den Spiegel 2001): solve the fallacy of large numbers with EUT</td>
<td>(Lopez 1987): SP/A Theory</td>
</tr>
<tr>
<td></td>
<td>(Shefrin and Statman 2000): Behavioural Portfolio Theory</td>
<td>(Lo 2004): Adaptive Market Hypothesis</td>
</tr>
</tbody>
</table>

Table 2: Different milestones in the paradigm of Behavioural Finance and the competing paradigm with efficient markets and rational agents. The puzzles are listed in bold and the background color of the theories is grey.

References


Bachelier, L. (1900a). Théorie de la spéculación. *Annales Scientifiques de l’école Normale Supérieure* 3(17), 21–86. 8, 49

Bachelier, L. (1900b). *Théorie de la spéculacion*. Paris: Gauthier-Villars. 8, 49, 74


Barber, B. M. and T. Odean (2001). Boys will be boys: Gender, confidence, and common stock investment. *Quarterly Journal of Economics* 116(1), 261–292. 68


Bernoulli, N. (1713b). formulation of the st peterburgs paradox. in a letter to Pierre Raymond de Montmart, problem 5, printed in *Essay d’Analysis* page 402. 5, 74


Cramer, G. (1728). Solution to the st. petersburg paradox. In a correspondence letter to Johann Bernoulli, cited by (Bernoulli 1738). 5, 6


De Brouwer, P. and F. Van den Spiegel (2001, Jan). The fallacy of large numbers revisited: the construction of a utility function that leads to the acceptance of two games while one is rejected. Journal of Asset Management 1(3), 257–266. 36, 71, 75


le Bon, G. (1896). *The Crowd, a study of the Popular Mind*. T. Fisher Unwin Ltd. 8

79
efficiency from an evolutionary perspective. *Journal of Portfolio Man-
agement* 30, 15–29. 73, 75
not follow random walks: Evidence from a simple specification test.
*Advances in Experimental Social Psychology* 21, 255–295. 38, 75
Mackay, C. (1841). *Memoirs of extraordinary Popular Delusions and the
Madness of Crowds* (First ed.). New Burlington Street, London, UK:
Richard Bentley. 8, 74
91. 29, 31, 74
Markowitz, H. A. (1952b). The utility of wealth. *Journal of Political Eco-
omy* 60, 151–158. 34, 35, 36, 74
Markowitz, H. A. (1959). *Portfolio Selection: Efficient Diversification of
Investments*. New York: John Wiley & Sons. 40
view* 50, 370–396. 23, 24, 56, 74
Penguin Compass. 22, 23
May, R. M. (1976, Jun). Simple mathematical models with very compli-
Menger, K. (1934, Aug). Das unsicherheitsmoment in der wertlehre betrach-
tungen im anschlu an das sogenannte petersburger spiel. *Zeitschrift für Nationalökonomie* 5. 7
Merton, R. C. (1969, Aug). Lifetime portfolio selection under uncer-
*Econometrica* 41(5), 867–887. 53


Pascal, B. (1654). Expected utility theory. in correspondence with Louis de Fermat on the subject of gambling. 74


82


Treynor, J. L. (1961). Market value, time, and risk. unpublished paper. 43, 74


83

