2021-2022

Masters, M2

Law and Finance

The Nature of Risk

Joseph Noko

Under the direction of Prof. Guido Hülsmann

Jury

Prof. Guido Hülsmann

Dr. Enareta Kurtbegu

École Supérieure d'Économie
et de Management des Patrimoines
Université d'Angers

CAUTION

The university does not intend to give any approval or disapproval to the opinions expressed in the work of the students: these opinions must be considered as their own.

Non-Plagiarism Commitment

Commitment to sign and attach to all reports, files, dissertations or thesis

I, the undersigned

Declare that I am fully aware that any plagiarism of documents or part of a document published on any form of support, digital or paper, constitutes a violation of copyright as well as fraud. Accordingly, I undertake to cite all the sources I have used to write this thesis / report / dissertation.

Signed by the student on 6/ 10 / 2022

Creative Commons License

Thanks

Professor Guido Hülsmann for his patience and inviting me to think more deeply about ideas, and to consult the great works of finance, Dr. Enareta Kurtbegu for sharing the Saint Petersburg Paradox and the delights of risk theory with me, Socrates for urging me to question the nature of things, and my friends and family for aeons of undeserved support. Thanks also to the positive provocations and inspiration from all-too fleeting encounters with Professors Catherine Defains-Crapsky, and Dushica Stevchevska-Srbinoska.

Contents

Non-Plagiarism Commitment	**2**
Thanks	**4**
Contents	**5**
Abstract	**7**
Introduction	**9**
Risk and Its Measurability	9
Risk and the Possibility of Loss	12
Methodology	**17**
Part I: On the Measurability of Risk and the Meaning of Probability	**18**
Franklin’s Conjecture: Truth and Probability Before Pascal	18
Rhetoric and the Search for Truth	19
Aristotelian Logic	20
Laws of Evidence in the Development of Probability	23
Probability and Ethics	25
The Price of Peril	27
The Pursuit of Scientific Truth	31
Verification of a consequence	31
The Duhem-Quine Thesis	32
Monotonicity	33
Scaling Arguments	33
Statistical syllogisms	34
Combining Arguments	35
Statistical Evidence	35
Frank Knight’s Theory of Risk and Uncertainty	35
A Theory of Knowledge	35
Subjectivity of Knowledge	36
On Error in the Estimation of Risk	37
Knight on Probability Situations	39
Indeterminateness and Knowability	41
The Limits of Knight’s Theory of Probability Situations	42
Von Mises’ Theory of Probability	43
Von Mises’ Scepticism About the Use of Mathematics	43
Time and Risk	44
Defining Praxeology	45
Reflexivity, Indeterminacy, and Uncertainty	45
On Case and Class Probabilities	47
Von Mises and the Frequentist Interpretation of Probability	49
Risk as an Epistemic Category	51
Keynes Theory of Probability	52
Keynes and the Logical Foundations of Probability	53
The Weight of an Argument	53

Keynes on Subjectivity and the Unquantifiability of Risks 54
Measurable Probabilities and Words of Estimative Probability 55
The Logic of Partial Belief 58
Keynes Contra Induction 59
The Art of Plausible Inference 60
Part II: The Prospect of Ruin 65
The Pascalian Revolution 65
The Maximisation of Value 65
The Problem of Points 65
The Weaknesses of Expected Value Maximisation 69
Risk Aversion, Situation Value and the Threat of Ruin 74
The Saint Petersburg Paradox 75
Daniel Bernoulli's Theory of the Maximisation of Geometric Returns 78
The Subjective Basis of Risk 78
Utility Derives from Growth 79
Bernoulli's Model 80
Nature's Admonishment to Avoid the Dice 85
Diversification and The Equity Premium Puzzle 88
Deconstructing the Equity Premium Puzzle 89
Caius the Merchant Discovers that Insurance Increases Wealth 90
Bernoulli's Error 92
Bernoulli's Heirs and Investing Under the Prospect of Ruin 93
Chaude Shannon and Choice, Uncertainty and Entropy 93
John Larry Kelly, Jr. and the Kelly Criterion 99
Conclusion 105
Appendix 108
Appendix 1: Data for the Capped Saint Petersburg Wager 108
Bibliography 110
Table of Figures 124
Table of Tables 124

Abstract

This treatise presents a pure theory of the nature of risk, arguing that aspects of risks are not quantifiable, that risk is subjective, and that losses impact portfolios more strongly than gains, leading to rationally obligatory loss aversion. Risk is built out as being involved not simply with probability distributions, but the weighing of evidence and judgement between ideas. This is done by reimagining probability as an extension of logic. Given the multiplicative nature of wealth, loss aversion is not only necessary to avoid gambler's ruin, but to grow it as well. The geometric mean is, therefore, the more appropriate decision criterion to build wealth and avoid the assumption of debilitating risks.

Keywords: measurability, uncertainty, truth, pre-Pascalian probability theory, subjectivity, Bayesian, logical probability, downside risk, losses, loss aversion

“It is a riddle, wrapped in a mystery, inside an enigma; but perhaps there is a key.”

Winston Churchill

INTRODUCTION

> "... Uncertainty must be taken in a sense radically distinct from the familiar notion of Risk, from which it has never been properly separated. The term 'risk,' as loosely used in everyday speech and in economic discussion, really covers two things which, functionally at least, in their causal relations to the phenomena of economic organisation, are categorically different. ... The essential fact is that 'risk' means in some cases a quantity susceptible of measurement, while at other times it is something distinctly not of this character; and there are far-reaching and crucial differences in the bearings of the phenomenon depending on which of the two is really present and operating. ... It will appear that a measurable uncertainty, or "risk" proper, as we shall use the term, is so far different from an unmeasurable one that it is not in effect an uncertainty at all. We ... accordingly restrict the term "uncertainty" to cases of the non-quantitative type.:" (Knight 1964, p. 19)

Since Frank H. Knight (1964) formulation of risk and uncertainty, his view of risk and uncertainty has become canonical (Runde 1998), influencing a range of eminent economists such as John Maynard Keynes (1937; 2013), Armen A. Alchian (1950) and G.L.S. Shackle (1968). Today, any discussion of risk follows the same lines, characterising it as defined by knowable probability distributions, and unknown specific outcomes, and uncertainty as being non-quantifiable, with an unknown probability distribution. Risk is a computational question, and "Knightian uncertainty" is what keeps central bankers and investors up at night.

Despite Knight's contribution, there has been a lack of clarity about the nature of risk, a murkiness in pure theory that extends to the Pascalian revolution, when the mathematisation of risk muddied the clear waters of pre-Pascalian probability theory. Not only is the risk literature replete with different definitions of risk, but, definitions are often applied inconsistently and with contradictory results (Crowe and Ronald C. 1967). Nevertheless, we can see two main features of risk, according to the Knightian consensus: risk is measurable, and it is a forecasting problem. This thesis will investigate those two features, with the first part dedicated to the question of risk's measurability, and the second, to the forecasting problem.

Risk and Its Measurability

While the work of forecasters often resembles blindfold, dart-throwing chimpanzees, (Malkiel 2015; Tetlock 2017; Tetlock and Gardner 2015), human action relies upon our ability to plan for a future of which we are uncertain, estimate what decisions will be most profitable, and to make decisions the fairness or appropriateness of which we cannot be sure (von Mises 1998; Knight 1964). Decision-making always occurs behind a veil of risk and uncertainty. The problem of risk and uncertainty is, fundamentally, an epistemic problem (von Mises 1998; Knight 1964; Franklin 2015). A decision maker, when thinking about risk and uncertainty, is, in effect, asking the question, “What do I know and how confident can I be about what I know?”. As epistemic categories, humanity has grappled with risk and uncertainty since the dawn of civilization (Franklin 2015). Yet, risk theory tends to assume that it is a very modern subject, that can be dated to the mathematisation of probability theory in the seventeenth century, from whence, risk has been viewed as a measurable, enumerable concept, and its analysis has proceeded along purely statistical lines (Bernstein 1996; Franklin 2015; Hacking 2006; Gabbay, Woods, and Hartmann 2011), culminating in a consensus around Knight’s view of risk. To modern theorists, the Pascalian Revolution is the moment when risk was born after millenia in which humanity knew only uncertainty.

The view that risk is measurable, rests upon the assumption that risk is objective. Pierre-Simone Laplace (1995) consolidated the post-Pascalian classical view of probability, which was grounded on a theory of the probability of events of the same class as being equiprobable, as a result of its roots in the problem of points. That view took probability as a feature of the physical properties of the world. The classical view could only handle a finite number of possible outcomes and an *a priori* determination of the equiprobability of events (Spanos 1986).

The view that probability is objective, has been cemented in the last century, with the emergence of the frequentist interpretation of probability, or aleatory probability, in which the probability of some outcome is the limit of its relative frequency of occurrence over repeated trials under similar conditions (Stark and Freedman 2016). Probability is, then, a function of the physical properties of the phenomenon being studied. This view can be traced back to the problem of points, which led Blaise Pascal (1665) to develop the idea of expected value. In order to determine the relative frequency of an outcome, it should be possible to have a large number of trials in order to smooth away the randomness of physical phenomena, and determine the underlying order. However, even with this interpretation of probability, it is possible to argue about the objectivity of the relative frequency: for example, the value judgments about the number of trials can lead to sometimes materially different relative frequencies. The relative frequency is, thus, not fixed, it is always at least somewhat inexact, in other words, we are forced to accept an error rate with every risk measure. Uncertainty is a feature of risk, not a separate category. In addition, as P.B. Stark and D.A. Freedman

(2016) argue, frequency probabilities are somewhat circular and subjective: for instance, it is not possible to actually have a large number of trials to forecast the relative frequency of earthquakes within the next thirty years. Those “trials” are not laboratory experiments. This is even more important when you consider that earthquakes are rare, unlike changes in weather, with recurrence over periods of hundreds of years. They argue that probability should be seen as a property of a mathematical model that seeks to describe features of the natural world, rather than an expression of the natural world. In doing so, once again, we are forced to accept that risk cannot be segregated so easily from uncertainty, because no mathematical model can perfectly express the natural world (Chaitin 2004). We are also forced to accept that many choices made within a mathematical model cannot be tested and may even seem arbitrary (Stark and Freedman 2016). While uncertainty estimates attempt to capture the error rate, they may very well fail to capture a number of materially important errors. Finally, in order to solve the convergence problem of determining when the outcome of a number of trials will start to approach a limit, we are forced into a circular definition of probability, in which probability is defined, given some uncertainty measure, which depends upon some factors, and that probability emerges only after a certain number of trials. (Stark and Freedman 2016). Interpreting probabilities is riddled with uncertainty, and suggests that some aspects of risk cannot be measured.

An exploration of pre-Pascalian probability theory through the lens of James Franklin (1994; 2015) allows us to re-examine the objectivity and measurability of risk. This discussion is further enriched by deepening our study of Knight’s work, and reviewing the work of economists Ludwig von Mises (1998), and Keynes (1937; 2013). While dramatically different in their broader economic theories, they shared a number of views regarding risk and uncertainty. They understood risk as an epistemic problem, questioned the quantifiability, not merely of uncertainty, but of risk, and they are in harmony with probability theorists such as Bruno de Finetti (1937; 1972; 2017), E.T. Jaynes (2003), George Pólya (1968), Abraham Wald (1939; 195), and Thomas Bayes (1763) in viewing risk as a subjective category. By subjective, one does not mean to say that probability is a creature of human caprice. All thinking must obey the rules of logic, in that sense, it is logical. Indeed, Keynes’ assumed that logical probabilities are, in this sense, objective. By subjective, one means that probabilities are not a property of the physical world, and certainly not of the world of human affairs. Bernoulli (1954), deepens this sense by also showing that risk is dependent on prior wealth, and where one sits on a trade. Thus, it is not correct to speak of risk as if it exists outside of us and as if one thing is risky or not risky for everyone.

The logical interpretation of probability addresses those aspects of probability that are completely divorced from the physical world. The modelling decisions that a mathematician makes, the decision to use one risk measure over another, or the credence given to one theory over another, are all examples of areas in which we are forced to weigh evidence without recourse to some final test that tells us what is true, or more likely. These logical, epistemic or inductive probabilities are a component of probability and define the unquantifiable risks that managers face. The nature of these probabilities is that they lead to only partial entailment or partial belief, which is to say, uncertainty is implicated in their formation. Thus, we are forced to embrace uncertainty as a component of risk, and the immeasurability of aspects of risk.

Part I of this treatise deals with the measurability of risk and the meaning of probability. Probability theory is, essentially, a means to understanding and predicting risk. Whereas the Pascalian Revolution set in motion a wave whose crescent was Knight's delineation of risk as, by necessity, a measurable thing, that is, something to which we can assign a number, probability theory in the pre-Pascalian world wrested on logical foundations, and was largely non-mathematical. Although there is a tendency to think that the Pascalian Revolution is a perfect successor of pre-Pascalian probability theory, accepting the baton and forging ahead, the two periods dealt with very different notions of risk. This is important even in the post-Pascalian world because both classes of risk exist to this day. For instance, investors have a smorgasbord of risk measures to use, but the risk of selecting the wrong risk measure cannot be measured. A theory, concept or argument may be wrong, even if the underlying data is perfect, and there is no way to arrive at a probability that one argument is more likely than another (Franklin 2015). The pre-Pascalian view of probability dealt with issues of weighing arguments, and theories, something which is outside the scope of the mathematical theory of probability. The older tradition of probability sees it as an extension of logic (Aristotle 2007; 350 AD; Arnauld and Nicole 1996; Jaynes 2003; Pólya 1968) which, in turn, leads to a fuller appreciation of the risks that investors face, and a larger sense of the tools investors need to assess risk.

Risk and the Possibility of Loss

It can be said that Knight does not so much as define risk and uncertainty, as state whether they can be measured or not, but in doing so, he opens the door to what are known as "downside" and "upside" risk. This is not unique to Knight: for example, Harry Markowitz (1952) equated risk with "volatility of returns" and this broad view of risk forms part of modern portfolio theory. The International Organisation for Standardisation (2009) defines risk as “the effect of uncertainty on objectives”, which again allows for a notion of risk that encompasses the good and the bad. This notion of risk, however, goes against our intuition. One does not, for instance, say, “There is a risk I will make a profit on this investment”. This notion of risk is also self-contradictory: it would seem perverse to imagine as rational behaviour a situation in which a person sought to limit all their risk, that is, both downside *and* upside risk. Defenders of the position would claim that when they refer to risk, they are of course referring to downside risk, which begs the question why the scope of “risk” sometimes allows for so-called “upside risk” to be defined as such. It seems evident that whereas an economic agent would want to limit their downside risk, they would be happy to have their profits run far in excess of what they expected[1]. The idea of risk as composed of upside risk and downside risk is a very agnostic view of risk, like a man who does not know if he believes in God and is not willing to say either way.

It should be clear that one views downside risk as the only legitimate risk. Indeed, Markowitz (1959) argues that his earlier use of variance as opposed to semi-variance was due to a lack of computing power and the greater familiarity that practitioners had with variance as opposed to semi-variance. The great man was aware of the inherent absurdity of so catholic a notion of risk, saying that semi-variance was a “more plausible measure of risk”. Nevertheless, this notion of risk is pervasive, partly because Markowitz himself continued to advance variance as a measure of risk, arguing that practitioners needed to become more familiar with the simpler measure of variance, before they could advance to the more plausible notion of risk.

A.D. Roy (1952) realised that an investor is not interested in what happens on average, but will happen in a particular instance, and that the maximisation of expected value is incompatible with diversification. Roy (1952, 432) criticised economic theory as being “set against a background of ease and safety”, rather than “poorly chartered waters” or hostile jungles, assumes “economic survival”, and thus being incapable of understanding why

[1] If, for example, a person is presented with two choices: the first, an option to invest in an instrument with a guaranteed 5% return and no higher, and the second, an option with a guaranteed 5% return with the potential for a higher return, it is evident that a rational economic agent would choose the latter instrument. It is only with the possibility of achieving less than the 5% return in the latter instrument, that the former instrument becomes more attractive, leading to the phenomenon of risk aversion, the preference of more certain outcomes over less certain ones. Thus, an economic agent would welcome “upside risk”, but frown upon “downside risk”. It is the possibility of loss that drives economic agents toward certainty, and, as one shall argue, that defines risk and leads to rationally obligatory loss aversion.

investors act as if they see disaster everywhere. Consequently, he modelled downside risk, or mean semivariance, because he felt that investors are driven by a “safety first” logic. James C.T. Mao (1970) surveyed managers and discovered a disparity between theory and practice. Across eight industries, managers said they believed that semi-variance was a more plausible measure of risk than variance. Frank Sortino (2010) developed post-modern portfolio theory, taking a myopic view of risk, focusing solely on downside risk, with an eye to deviations from target returns.

It shall be argued that risk should be viewed myopically, as being primarily the possibility of an economic agent incurring[2] a loss, and secondarily, the possibility of underperforming some target rate of return. By loss is meant a decay or decline in prior wealth, where wealth represents the assets of an economic agent, a part of which is risked in some venture, or risky proposition.

The myopic focus on risk becomes even more important when we consider the changes in behavioural economics in recent decades. Daniel Kahneman and Tversky (1979) discovered that people are more fearful of losses than they are desirous of making corresponding gains, something they called, “loss aversion”. Loss aversion is often treated as irrational, and managers have been encouraged to overcome their loss aversion in search of superior returns (Damodaran 2014; Koller, Goedhart, and Wessels 2020). However, if risk is defined not just by the question of its measurability, but also by the possibility of loss, then loss aversion is a rationally obligatory response, whereas, if risk is merely a forecasting error about a range of outcomes, good and bad, then, as prospect theory argues, and as managers are taught to believe, loss aversion is an irrational bias that should be overcome.

Part II focuses on the thoughts of Daniel Bernoulli (1738), Claude Shannon (1971), and John L. Kelly, Jr. (1956). They represent a tradition that takes a myopic view of risk, one defined by an allergy to loss, based on the multiplicative nature of wealth[3]. Bernoulli is an important figure in this regard. Although he is seen as the father of expected utility theory, inspiring John von Neumann and Oskar Morgenstern (1953)’s work, a closer reading of his work

---

[2] Following Crowe and Ronald C. (1967), I specifically use “incurring” to highlight that knowledge of loss is not essential, what counts is the moment and existence of loss.

[3] By this is meant, quite simply, that wealth accumulation is defined by a multiplicative function. So, for instance, year-over-year growth in wealth is due to a certain percentile growth in wealth in that period. A crude example will demonstrate this; suppose one starts the year with a fortune of €1 million and it grows to €1.5 million, over the next year. That can be expressed as €1 million*1.50 = € 1.5 million. It is for this reason that the ultimate number by which investors are judged is their rate of return.

shows that it speaks to the nature of risk, and the importance of avoiding loss in the accumulation of wealth. Indeed, rather than suggesting expected utility as a decision criterion, Bernoulli seems to use it as a pedagogical device to explain his findings, and propose the use of a geometric mean as a decision criterion for investors. Bernoulli shows that risk is subjective, depending on one's prior wealth. This idea can be stretched to show that risk for one is also an opportunity for another, providing the basis of mutually beneficial exchange. Bernoulli shows that wealth is multiplicative. The multiplicative nature of wealth leads naturally to the conclusions that in reducing risk, we increase our possibility of building wealth. From Bernoulli, we shall see that losses impact portfolios more profoundly than corresponding gains. In truth, although perhaps controversial, this can be shown with very elementary arithmetic: a decay in wealth of 10% requires an 11.11% gain for an investor to break even, while a 20% decline requires a 25% gain, a 50% decline demands a 100% gain, and a 99% decline requires a miracle. As losses mount, the gains needed just to break even escalate asymmetrically. Our discussion will show the logical basis of this, suggesting that, given the concavity of the logarithm of wealth, investors are rationally obliged to avoid risk, investing only when they have an edge, and, when doing so, investing in concentrated portfolios, while diversifying across time. In making this argument, we shall establish the maximisation of the geometric mean as a rationally obligatory criterion for investing in a world where losses are more impactful than gains.

Bernoulli's work is perhaps the earliest attempt at understanding non-ergodic processes. Recent work on ergodicity (Peters 2011; 2017; 2019; 2020; Peters and Adamou 2018; Peters and Gell-Mann 2016) has been largely ignored within economics. However, Bernoulli's insight, and the development of Claude Shannon's (1971) information theory, and John L. Kelly, Jr. 's (1956) betting and investing criterion, show that ergodicity is an important concept for understanding the nature of wealth, and risk. Shannon and Kelly echo Bernoulli in arguing that risk-seeking behaviour is at odds with the pursuit of wealth, and that, absent an edge, investing is *irrational*, because it destroys wealth. Bernoulli (1954, 29) calls this "nature's admonishment to avoid the dice altogether". Given that we are obliged to invest to build wealth, Bernoulli provides the world's first mathematical proof of diversification, centuries before Markowitz (1952). Bernoulli also argued that insurance is necessary to build wealth. From Bernoulli, we can rationalise the existence of the equity premium as being because of the risk of wealth destruction that the typical stock carries. When an investor has an edge, Bernoulli and Kelly ask us to invest heavily, in concentrated portfolios.

Methodology

Joan Robinson (1932, 6) said, "the two questions to be asked of a set of assumptions in economics are: Are they tractable? and: Do they correspond with the real world?", going on to say that, "...more often one set will be manageable and the other realistic," and that some branches of theory may be both manageable and realistic. In the treatise that follows, we shall question the assumptions surrounding risk, and in doing so, one hopes to develop a theory of the nature of risk that is both manageable and realistic, showing the rational basis of commonsense notions of risk.

This is a work of pure theory and the approach taken here is to give an exposition of the main ideas on pre-Pascalian probability theory, from which extensive use of James Franklin's (2015; 1994) work, and deconstruct, the meaning and coherence of these ideas, and to build upon what insights are gained, by interrogating the main ideas on probability from Knight (1964), who gave us the present consensus on risk, von Mises (1998), and Keynes (2013a; 2013b). The second part tackles the relationship between risk and loss, largely through a re-reading of Bernoulli (1954), which is really about the nature of risk. By re-engaging Bernoulli, we can enrich our understanding of the nature of risk. Finally, we shall also examine the work of his heirs, Claude Shannon (1971), the father of information theory, Edward Thorp (1969; 1966; 1971; 1997), and John Kelly, Jr. (1956) to deepen our understanding of Bernoulli's contribution, and to suggest an answer that Bernoulli was not quite able to answer: how to invest given his main finding, that losses impact gains more profoundly than corresponding gains.

Ideas will be examined from both a logical point of view, and, where appropriate, a statistical point of view, to understand the implications of an idea. Given the nature of the question at hand, little attention will be given to decision criterions and theories of decision making under uncertainty. While some attention will be paid to expected value theory, this will be as a way of explaining the emergence of the Saint Petersburg paradox, and as a way of introducing Bernoulli's theory of risk.

Part I: On the Measurability of Risk and the Meaning of Probability

Franklin's Conjecture: Truth and Probability Before Pascal

When, in 1736, Joseph Butler said that "Probability is the very guide of life," he was pointing to the development of the mathematical theory of probability as a guide to decision making (Franklin 2015). He was also asserting the reality that decisions are largely made in the absence of certain knowledge. Since the mathematisation of risk theory during the Pascalian Revolution, there has been a certain tendency toward imagining that risk theory began in the seventeenth century, with man crossing the line from uncertainty to risk, from a world in which the future was the plaything of gods, to one whose outlines could at least be calculated (Bernstein 1996). Ian Hacking (2006) goes so far as to say there was no probability theory prior to the mid-seventeenth century, while other historians, such as Peter L. Bernstein (1996) and Dov M. Gabbay, John Woods, and Stephan Hartmann (2011) have merely skirted over pre-Pascalian risk theory.

Yet, prior to Blaise Pascal and Pierre de Fermat's discovery of the mathematical theory of probability in 1654, man still had need of tools for navigating through the murky future and assessing commercial risk. Indeed, man has successfully navigated thousands of years of history without the aid of the mathematics of probability. Furthermore, the main ideas around the weighing of uncertain evidence and assessing commercial risk were developed in ancient times, and by Scholastics and legal writers in the late middle ages (Hannam 2010; Franklin 2015).

James Franklin's *The Science of Conjecture* is perhaps the only book that treats pre-Pascalian ideas about probability and decision making under uncertainty. His history of pre-Pascalian probability theory provides a foundation upon which we can reimagine the nature of probability and, by consequence, the nature of risk. It must be said that Hacking (2004) argues that even if Franklin is right and that there was a tradition of probability theory before its mathematisation, the mathematical theory of probability is its rightful successor, and the ancient tradition is of little real use. Franklin's history of this neglected period tells us that the history of the pre-Pascalian world is not a tale of luck, with hapless, helpless humanity braving thousands of years of history as they awaited the mathematization of probability theory. Hacking suggests that this history, while enriching, is of no material significance to modern probability theory. However, a review of the main lines of pre-Pascalian probability theory, will show that that modern notions of risk are only one way of viewing risk, and that we might gain from recovering this older tradition, where probability theory, in its guise as logical or epistemic probability, tackled broader questions of decision making under uncertainty, in courtrooms, in philosophical debates, on the high seas and a vast panorama of other settings. By studying the pre-Pascalian theory of probability, we can arrive at a broader view of probability theory that will in turn allow us to reframe risk theory and enlarge the tools that managers and investors have, for mitigating risk.

Franklin (2015) asserts that there are three kinds of probability: unconscious inferences that are studied by psychology; the mathematical theory of probability, which emerged with Pascal and Fermat; and “ordinary language reasoning” about the plausibility of things, which dominated probability theory until the Pascalian Revolution. Ordinary language reasoning about probabilities need not make use of numbers, as with “proof beyond a reasonable doubt” in law, or, instances of plausible reasoning made by scientists and policeman, or, if it does make use of numbers, it is in crude forms typical of guesstimations (Franklin 2015; Jaynes 2003; Bernstein 1996). The history of pre-Pascalian probability theory is the history of ordinary language reasoning about probabilities. Franklin (2015) reminds us that the mathematical theory of probability and ordinary language reasoning map onto two types of probability: factual, or stochastic or aleatory probability, which governs chance events that lead to random sequences, and logical, or epistemic or non-deductive probability, which deals with the weighing of evidence between propositions. The pre-Pascalian tradition is, therefore, a tradition grounded in the logical interpretation of probability, evolving along four main lines: the rhetorical tradition, laws of evidence, the evaluation of commercial risk, the rules of scientific enquiry.

Rhetoric and the Search for Truth

In ancient Greek times, probability theory was seen as the domain of rhetoricians, and used to investigate a wider range of arguments than is possible with a strictly mathematical view of probability. Although the rhetorical tradition was developed for use in courts of law, it leaned heavily towards persuasion. Indeed, it can be said that because it tended to conflate the search for the most logical argument with the desire to find the most persuasive argument, it was unable to foster an independent theory of non-demonstrative argument (Franklin 2015).

Socrates, who defended himself against accusations of corrupting Athenian youth by resorting to a probabilistic defence of his innocence, defined the Greek word for probability, εικός[4] (*eikos)*, as “likeness to truth”, similar to our notion of plausibility (Franklin 2015; Bernstein 1996). Likeness to truth, of course, suggests very clearly that one’s conclusions are not necessarily the truth. In the absence of certain knowledge, one makes a plausible inference. The Socratic dialectic is an example of an ideal mode of plausible reasoning in pursuit of the truth (Bernstein 1996). Logical probability was not about experimentation and quantification, but working through a logical chain of reasoning toward a plausible conclusion. To the Greeks, what counted was one’s ability to prove something with logic and axioms. Bernstein (1996) observes that, although Archimedes claimed to be able to move the world with a lever, he never bothered to try. The Greeks were not a people wont to experiment, and under them, the growth of a mathematical theory of probability was stifled, but probability as a rhetorical tool in the search for truth, flourished.

The Greek word *πιθανός*[5] (*pithanós*) takes on a range of meaning from the merely persuasive to the plausible (Franklin 2015). For instance, when Herodotus says that a certain story is “the most plausible of the many stories of Cyrus’ death”, he uses the word *πιθανός* (Franklin 2015, 103). The Greek tradition also sought to “hear both sides” of an argument, in the quest for the truth of a matter, a harbinger of modern court systems. Although the Greek rhetorical tradition, at least in the fifth century with the flowering of rationalism, had the seeds for a non-demonstrative theory of argument, the success of the Sophists in helping lawyers win cases, overwhelmed the search for truth. The Sophists cared about persuasiveness, not plausibility.

Aristotelian Logic

Aristotle’s seminal text *On Rhetoric*, is the most important work on probability theory of the pre-Pascalian era (Franklin 2015). Although Aristotle sees rhetoric as being preoccupied with *πιθανός*, he divides arguments into two kinds: logical arguments, which are the subject of dialectic, and sophisms, which are the subject of rhetoric. Aristotle (2007, 30) explains the role of each saying,

[4] eikos.

[5] Pithanon.

> Rhetoric is an *antistrophos*[6] to dialectic; for both are concerned with such things as are, to a certain extent, within the knowledge of all people and belong to no separately defined science. A result is that all people, in some way, share in both; for all, up to a point, try both to test and uphold an argument [as in dialectic] and to defend themselves and attack [others, as in rhetoric].

Aristotle believed that deductive arguments were bounded by their inability to arrive at a firm conclusion about things. Instead, a person must call upon *εἰκότος* (likelihoods) and and signs, in the search for the most plausible proposition, *εικός* (Aristotle 2007; Franklin 2015). In terms of probability, Aristotle drew upon induction, or *ἐπαγωγή*[7] (*epagōgḗ*), arguments from example, signs, or *σημεῖᾰ* (*sēmeîa*) and analogy (Aristotle 2007; Franklin 2015), saying,

> In the case of persuasion through proving or seeming to prove something, just as in dialectic there is on the one hand induction [*ἐπαγωγή*] and on the other the syllogism and the apparent syllogism, so the situation is similar in rhetoric; for the *παραδειγμα* (*paradeigma)*[8] is an induction, the *ἐνθύμημα* (*enthýmēma)* a syllogism.

Εικός and *σημεῖᾰ* refer to human action in different ways. Whereas *εικός* corresponds to "likeness to truth", for Aristotle, *σημεῖᾰ* corresponds to things that are necessarily true, with some signs being necessary, and Aristotle referred to these as *τεκμήριον*[9] (*ekmḗrion*) and others merely probable, fallible signs, or *σημήϊον* (*sēmḗïon*). When a person says something that they think is irrefutable, for instance, "He is angry because something has upset him", they think they are providing a *τεκμήριον*.

Aristotle (2007, 190-191) understood his framework as being a logical one rather than a demonstrative one:

> …since the probable is not always but for the most part true, it is clear that all these kinds of *enthymemes* can be refuted by bringing objections, but the refutation may be an apparent one, not always valid; for the objector does not refute the probability but [shows] that [the conclusion in a particular case] is not necessary. As a further result,

[6] Antistrophos means "counterpart", or even "correlative" and "coordinate." It can also mean "converse."

[7] Literally, leading in, bringing in or upon.

[8] Pattern, example, or sample.

[9] Proof through argument, or demonstrative proof.

> the defendant always has an advantage over the prosecutor because of this fallacy; for since the prosecutor demonstrates by probabilities, and since it is not the same thing to show in refutation that an argument is not probable as to show it is not necessary, what is for the most part true is always open to objection; for otherwise it would not be for the most part and probable, but always and necessary. Yet when a refutation is made in this way, the judge thinks either that the thing is not probable or that it is not for him to decide, reasoning falsely as we said—[falsely] because he should not only judge from necessary arguments but from probable ones, too; for this is to judge "in accordance with the best understanding." Therefore, it is not sufficient if one attacks an argument as "not necessary"; and one must refute it as "not probable." This will result if the objection is a more probable statement; and that such is the case can be shown in two ways, either in terms of the time or the facts but most effectively if by both; for if something is often true, it is more probable.

Aristotle's treatment of rhetoric was potentially the basis of a theory of probability, but as we have said before, Greek civilization was unable to build on that (Sambursky 2014; Bernstein 1996; Franklin 2015). This is all the more remarkable when one considers that the Greeks, and indeed, the Romans, were obsessed with games of chance. Their randomness fascinated ancient man. Indeed, games of chance were played across the ancient world, with the first games of chance likely played in Egypt, with knuckle bones, or *astragali* (Sambursky 2014; Bernstein 1996). The earliest depictions can be found in Egyptian tomb paintings from 3,500 B.C. (Bernstein 1996). These ancient games of dice were the progenitors of what Arabs called, الزهر (*al-zahr*), for dice, which eventually became known as games of "hazard". Yet, the Greeks did not seek to build a theory of risk based on games. Whereas the mathematicians of the sixteenth and seventeenth century took them as a challenge, for the Greeks, this was the domain of the gods (Sambursky 2014; Bernstein 1996; Franklin 2015). In fact, the ancient world was unable to make even the most minute advance toward observing any regularities in random series.

Shmuel Sambursky (2014) suggested that Greek scientific progress was limited because of the Greek conception of fate, which prevented them from perceiving any deeper order beneath the randomness of life. Greek tragedy, after all, was a tale of people who, regardless of their decisions, were fated to some end. The Greek conception of fate did not allow for a belief that one could project oneself into the future and imagine various courses of action. Bernstein (1996, 18), quotes the Egyptologist, Henri Frankfort, "The past and the future -far from being a matter of concern- were wholly implicit in the present." Furthermore, the Platonic distinction between truth and probability held that truth was to be found in pure forms, and Greeks in general were not predisposed to seek the truth in physical reality. This highly popular Platonic view resulted in Greeks being so persuaded by the power of deduction that they believed induction was largely unnecessary.

Further halting the progress of the development of a mathematical theory of probability is that Aristotle's *On Rhetoric* was largely unknown in Western Europe from the time of Quintilian until the third and fourteenth century. The Roman Empire and its successors experienced a decline in the quality of the rhetorical tradition, which, instead, passed on to the Islamic world, with Aristotle's work translated into Arabic (Franklin 2015). The Islamic world correctly understood Aristotle's *On Rhetoric* and his *Poetics* as part of his work on logic, and so, it gave them more credence than did the European thinkers of the Middle Ages. For instance, Al-Farabi is said to have read Aristotle's *On Rhetoric* two hundred times and written seventy books about it (Franklin 2015). Ibn Sina, known in the West as Avicenna, was profoundly influenced by Aristotle in writing his book, *Treatise on Logic*. Nevertheless, in Europe and the Islamic world, there was no development of a mathematical theory of probability from Aristotle's work.

Laws of Evidence in the Development of Probability

Law has had an immense influence on the theory of probability. Franklin (2015), having studied ancient Egyptian, Mesopotamian, Talmudic, Greek and Roman and mediaeval European texts, observes that lawyers have to work with a plethora of concepts which contest for applicability in specific cases, and often, it is not possible to come to a certain conclusion about a matter, rather, decisions are made on the basis of the opinions of lawyers, experts and juries. Not only has law always been a testing ground of plausible reasoning, but, the discoverers of mathematical probability were either lawyers, in the case of Fermat, Huygens, de Witt, and Leibniz, or sons of lawyers, in the case of Cardano and Pascal.

The importance of lawyers, or people who had studied law, in the development of mathematical probability, is due to the all-encompassing importance of law in ancient times to the Renaissance. A legal training was essential for success not simply in law but in the world of affairs outside medicine and theology (Bouwsma 1973; Franklin 2015). It was assumed that every educated person had some understanding of the law. Often, mathematical progress was driven by legal concerns. For instance, the development of arithmetic was driven by the need to accurately calculate legal obligations, such as taxes, debt, and profits, and lawyers required some progress in geometry to aid in the measurement of land (Maanen 1992; Franklin 2015).

Laws of evidence were methods steeped in fields such as law, science, commerce, philosophy, and logic and emerged out of Roman law (Franklin 2015). Leibniz believed the law of evidence was a logic of probability. It was a science of conjecture grounded in the use of ordinary language reasoning about probabilities. The nature of logical probability is that success is about sifting through evidence to determine the most plausible evidence. For instance, E.T. Jaynes (2003, 3) provides the following example,

> Suppose some dark night a policeman walks down a street, apparently deserted. Suddenly he hears a burglar alarm, looks across the street, and sees a jewellery store with a broken window. Then a gentleman wearing a mask comes crawling out through the broken window, carrying a bag which turns out to be full of expensive jewellery. The policeman doesn't hesitate at all in deciding that this gentleman is dishonest.

The example is not an instance of a logical deduction from the evidence, but it is an example of the use of plausible reasoning by way of weak syllogisms (*ἐπᾰγωγή,* or *epagōgḗ*), in this case, of the form:

$$\begin{array}{c} \text{If A is true, then B becomes more plausible} \\ \text{B is true} \\ \hline \text{therefore, A becomes more plausible} \end{array} \tag{1}$$

As a general guide, plausible reasoning uses prior information to determine the degree of plausibility of a proposition, and as aforementioned, this chain of reasoning is often unencumbered by any need for quantification. Law, and even a science such as medicine, have remained resistant to the Pascalian revolution, instead, proceeding along a different course of probability. Indeed, the field of probability until the Renaissance was dominated by Scholastics and lawyers, rather than mathematicians. Given the importance of Scholastics and lawyers to the development of what we can class as pre-modern probability theory, and therefore, pre-modern ideas of risk, the law of evidence was the most important principle of this long era, regulating what evidence was admissible, the grades of evidence, half-proofs[10], what sources were credible, how conflicting evidence was to be treated, and whether hearsay was proper (Franklin 2015). The Egyptians, Greeks, Romans, indeed, every society developed its own laws of evidence to help them determine the most plausible proposition, although, in some instances, where there was circumstantial evidence, the degree of certainty possible was lower than in other instances with more robust evidence.

Probability and Ethics

Although in modern times, one assumes that each person can come to the right conclusion on a moral question without the aid of some guide, this was not always the case. Aristotle first made the association between plausible reasoning and ethics, in his *Nicomachean Ethics*, where he remarked (Franklin 2015, 64),

> …it is the mark of an educated man to look for precision in each class of things just so far as the nature of the subject admits; it is evidently equally foolish to accept probable reasoning from a mathematician and to demand from a rhetorician demonstrative proofs.

An excellent example of the use of plausible reasoning in ethics is that between 1200 and 1650, Catholic casuistry guided people through “cases of conscience” (Franklin 2015). These guides began with the intentions of the individual, and provided rules to follow to assuage the doubting conscience. It is from these texts that we get the phrase, “moral certainty” and “probabilism”. In fact, probabilism in the seventeenth century referred to this doctrine in moral theology, a doctrine that fell under the weight of Pascal’s attacks in his *Provincial Letters*, establishing the modern notion that each man was his own fair judge (Franklin 2015).

Other ethical questions surrounded justifications for war, the morality of the state, and civil disobedience. For example, in the sixteenth century, Francisco Suarez, wrestled with the degree of certainty that a person must have of the evil of a law before they can disobey it. He concluded that without moral certainty, it was unethical to disobey the law of the land. He distinguished between negative doubts, in which there are no reasons for either proposition, and positive doubt, in which there are reasons for either proposition. Where there is negative

---

[10] In Latin, *semiplena probatio*. A concept from Roman law between mere suspicion and full proof (*plena probatio*).

doubt, the probability, say, of a soldier fighting in a war that he doubted was just, was higher, without any need to examine the issues. Where there was positive doubt, then the truth of the matter had to be excavated, and if that proved impossible, the more probable proposition must be followed. Pierre-Simon de Laplace and later Keynes both tackled the problem of insufficient reason, where on the one hand, there is a proposition with little evidence and a certain degree of probability, and on the other, one with lots of evidence and a similar degree of probability (Franklin 2015; Jaynes 2003). That problem remains unresolved to this day. For instance, while Knight (1964) believed that the odds of any face of a die, or side of a coin coming up are an *a priori* result of the physical properties of the die, or coin, and therefore, a reason to be indifferent toward the outcome, there are those who have argued that the many throws of coins and dice showed that there was a symmetry in outcomes.

Perhaps the most famous incident in this tradition of plausible reasoning is that of the Jansenists and their struggle with the Jesuits over the sufficiency of grace. It is arguable that Pascal's treatment of probability was a direct consequence of this *petite guerre*. Jesuits believed in free will, to such a point that they were faced with a simple question: if God gave two people identical graces, would they both be saved? Luis de Molina answered in the negative, enraging Dominicans and leading to a long examination by the Catholic church of the subject. Eventually, both views were deemed acceptable, with both sides instructed not to malign the other with accusations of heresy, and instead await a decision by the Holy See. That decision has never been made (Franklin 2015). The controversy flared up again when Cornelius Jansen, bishop of Ypres, published his book, *Augustinus*. In France, the Jansenists centred around a convent near Versailles, Port-Royal. As the Jansenists fought back, sustaining, until then, successful attacks by the Jesuits, Antoine Gombaud, chevalier de Méré, suggested a line of attack to his fellow Jansenist, Pascal. In the fifth of his *Provincial Letters*, Pascal lambasts the moral views of Jesuits, arguing that they stem from laxism. Pascal stated, incorrectly, that Jesuits believed that an opinion was probable, and therefore permissible, if at least one doctor held it. A person might, therefore, follow an opinion contrary to one's beliefs, if one "worthy and learned" doctor, as the Jesuit Thomas Sanchez termed him, affirmed an opinion. This laxism, which flung open the door to licentiousness, duly scandalised Pascal. For Pascal, laxism was grounded in probabilism:

> Voilà de quelle sorte ils se sont répandus par toute la terre à la faveur *de la doctrine des opinions probables*, qui est la source et la base de tout ce dérèglement[11]. (Pascal 2001, 51)

Pascal went on to say, with some exaggeration, that,

> *Une opinion est appelée probable, lorsqu'elle est fondée sur des raisons de quelque considération. D'où il arrive quelquefois qu'un seul docteur fort grave peut rendre une opinion probable.* Et en voici la raison *: car un homme adonné particulièrement à l'étude ne s'attacherait pas à une opinion, s'il n'y était attiré par une raison bonne et suffisante*. Et ainsi, lui dis-je, un seul docteur peut tourner les consciences et les bouleverser à son gré, et toujours en sûreté[12]. (Pascal 2001, 55-56)

Pascal's real gripe was that, unlike the Jesuits, he did not believe that there were any experts that a person could consult on moral issues. For him, a person should be their own fair judge . With that letter, there was a rupture between the ancient world and the modern, between a world where probabilism was used to assess moral issues, and one in which expert opinion and the use of probability were limited to matters of the world (Franklin 2015). Pascal greatly distorted the opinions of his enemies, but his work was transformative, with matters of conscience decided, not by authorities and their rationale, but by one's own conscience.

Ultimately, law has been a fertile ground for probability theory, and yet no legal system assigns a number to the likelihood that a case or evidence proves a charge beyond a reasonable doubt.

The Price of Peril

Although the pre-Pascalian world did not develop a mathematical theory of probability, it still made extensive use of probability, for the pricing of risk (Bernstein 1996; Franklin 2015).

---

[11] Translation: "This is how they have spread all over the world under cover of the doctrine of probable opinions, which is the source and the basis of all this disorder".

[12] Translation: "An opinion is called probable when it is based on reasons of some consideration. Whence it sometimes happens that a single very serious doctor can render an opinion probable. And here is the reason: for a man particularly devoted to study would not attach himself to an opinion, if he were not attracted to it by a good and sufficient reason. And thus, I told him, a single doctor can turn consciences and upset them as he pleases, and always in safety."

In ancient Greece and Rome, maritime transport was more efficient than land transport, even for medium distance trade. Ships could carry about 70 tonnes, and were run by a ship owner who carried freight for their customers, while also trading their own goods. The most costly capital item was the goods aboard the ship, for example, a cargo load of wheat was approximately the same value as that of the ship. The ship owners and merchants needed to raise funds to procure their goods, and given the unique risks of maritime trade, a special type of loan was devised for them, the *ναυτικὸν δάνειον* (*nautikòn dáneion*) or maritime loan[13]. The earliest known aleatory contract is the ancient Athenian maritime loan. The most informative description of this maritime loan is given by Demosthenes in the 4th-century BCE, in his speech, *Against Lacritus* (Millett 1983; De Ste Croix 1974; Franklin 2015; Hoover 1926; Cohen 1992). Demosthenes was involved in a case involving two merchants, Artemo and Apollodorus, from Phaselis, a city on the coast of ancient Lycia, in present day Türkiye. The men had been introduced to Androcles of Sphettus, and Nausicrates of Carystus, who lent them three thousand drachma in silver. Demosthenes submitted the agreement[14] to the

[13] In Rome, it was known as the *fenus nauticum*.

[14] The text of the agreement reads:

> Androcles of Sphettus and Nausicrates of Carystus lent to Artemo and Apollodorus, both of Phaselis, three thousand drachmae in silver for a voyage from Athens to Mende or Scione, and thence to Bosporus--or if they so choose, for a voyage to the left parts of the Pontus as far as the Borysthenes, and thence back to Athens, on interest at the rate of two hundred and twenty-five drachmae on the thousand; but, if they should sail out from Pontus to Hieron after the rising of Arcturus, at three hundred on the thousand, on the security of three thousand jars of wine of Mende, which shall be conveyed from Mende or Scione in the twenty-oared ship of which Hyblesius is owner.
>
> **[35.11]** They give these goods as security, owing no money upon them to any other person, nor will they make any additional loan upon this security; and they agree to bring back to Athens in the same vessel all the goods put on board in Pontus as a return cargo; and, if the goods are brought safe to Athens, the borrowers are to pay to the lenders the money due in accordance with the agreement within twenty days after they shall have arrived at Athens, without deduction save for such jettison as the passengers shall have made by common agreement, or for money paid to enemies; but without deduction for any other loss. And they shall deliver to the lenders in their entirety the goods offered as security to be under their absolute control until such time as they shall themselves have paid the money due in accordance with the agreement.

court, and it shows that Artemo and Apollodorus were to use a vessel owned by Athenippus of Halicarnassus (in western Türkiye) for “a voyage from Athens to Mende or Scione, and thence to Bosporus--or if they so choose, for a voyage to the left parts of the Pontus as far as the Borysthenes, and thence back to Athens”. They were to pay an interest rate of 22.5%, but if they sailed from “Pontus to Hieron after the rising of Arcturus [mid-September]”, that interest rate would climb to 30%, because this was the end of the safe sailing season. They were to buy three thousand jars of wine from Mende or Scione, which would act as security for the loan, and convey it to the Bosphorus, and use the revenue to buy goods as return cargo. The merchants claimed the ship had sunk and they were unable to repay the loan, as repayment occurred only if the ship arrived at its destination safely (Franklin 2015). The reasoning behind this was that it did not make sense to try and recover the loan amount from a person who had lost all their goods and quite possibly their own life. However, these loans were, in effect, insurance contracts, or more accurately, combinations of loan and insurance contracts, spreading risk, and granting some compensation to merchants in the event of disaster (Hoover 1926; Franklin 2015). This allowed funds to be loaned out without falling afoul of usury laws, and the borrower got what amounted to insurance against maritime disaster over the borrowed funds (Hoover 1926).

Already, we can see that risk was priced for different scenarios. In this case, the distance travelled and the safe sailing season affected the price of risk. The parties were assumed to

---

> **[35.12]** And, if they shall not pay it within the time agreed upon, it shall be lawful for the lenders to pledge the goods or even to sell them for such price as they can get; and if the proceeds fall short of the sum which the lenders should receive in accordance with the agreement, it shall be lawful for the lenders, whether severally or jointly, to collect the amount by proceeding against Artemo and Apollodorus, and against all their property whether on land or sea, wheresoever it may be, precisely as if judgement had been rendered against them and they had defaulted in payment.
>
> **[35.13]** And, if they do not enter Pontus, but remain in the Hellespont ten days after the rising of the dogstar, and disembark their goods at a port where the Athenians have no right of reprisals, and from thence complete their voyage to Athens, let them pay the interest written into the contract the year before. And if the vessel in which the goods shall be conveyed suffers aught beyond repair, but the security is saved, let whatever is saved be the joint property of the lenders. And in regard to these matters nothing shall have greater effect than the agreement.
>
> Witnesses: Phormio of Peiraeus, Cephisodotus of Boeotia, Heliodorus of Pitthus

understand the risks that they took on in their endeavour. The ancient Athenian maritime loan was embraced by the Romans and spread across the ancient Western world (Millett 1983; De Ste Croix 1974; Franklin 2015; Hoover 1926; Cohen 1992). Roman law was more explicit in its formulation of the maritime loan, and made the added assumption that risks were knowable and could be quantified. For instance, Roman stated that, “There can be no sale without a thing to be sold. . . . Sometimes, indeed, there is held to be a sale even without a thing, as where what is bought is, as it were, a chance (*quasi alea*). This is the case with the purchase of a catch of birds or fish or of largesse showered down. The contract is valid even if nothing results, because it is a purchase of an expectancy (*spei*)”, (Franklin 2015, 259-260). Roman law also spoke of “how much the hope of indebtedness is” (Franklin 2015, 260), which referred to the likelihood of repayment of a loan, and gave guidelines regarding the sale of an expected inheritance.

Under Roman law, the risk of a shipwreck was assumed by the state. Risk was reified and regulations defined under which risk could move from one party to another. Maritime loans were allowed to have higher interest rates than permitted for other loan products, because of the heightened uncertainty of maritime trade, with the *Digest* saying, “the price is for the peril” (Franklin 2015).

Maritime loans, and the underlying risk-sharing partnership between an entrepreneur and capitalist, were the only way of raising capital for such endeavours, such that the return encompassed the cost of risk, interest and foreign exchange (Hoover 1926; Franklin 2015). Under Roman law, the doctrine that, “he who bears the risk should get the benefit”, held, and the person bearing such risk included those who owned and used the goods placed under risk. This was to encourage capitalists to fund entrepreneurs. The fruitfulness of the maritime loan was such that not only did the Romans and much of the ancient Western world embrace it, but, around the twelfth century, it was revived along with maritime trade (Hoover 1926; Franklin 2015). Despite their commercial success, the price of peril skirted around usury laws. In fact, the interest rate charged for these loans were defined, not as interest rates *per se*, but as a bonus paid for the use of those funds, a legal fiction given grounds because repayment was not sought when a ship failed to escape maritime disaster (Hoover 1926). Care was taken in the drafting of contracts to avoid framing interest rates as such, and the term “usury” was avoided.

The Pursuit of Scientific Truth

Verification of a consequence

Verification of a consequence (Pólya 1968; Jaynes 2003; Franklin 2015) is the principle that each verification of an idea makes it more credible. George Pólya gives the example of Leonard Euler's conjecture that any number of the form 8n + 3, where n is positive, is the sum of a square and the double of a prime. That conjecture remains neither disproved or proved. Yet, when Pólya wrote, that conjecture had been proven to numbers under 1,000. Although each verification makes Euler's statement more credible, it does not lead to proof and Euler's conjecture remains open to falsification. The principle of verification of a consequence extends beyond proof theory and can be expressed as following the pattern,

$$\begin{array}{c} \text{if A is true, then B is true} \\ \text{B is true} \\ \hline \text{A is more plausible} \end{array} \tag{2}$$

It is well-nigh impossible to assign a number to the degree of confidence that one can have in such reasoning, even though it is a kind of reasoning that we use every day. The problem of induction and Humes' notion of a black swan is perhaps the most famous instance of the difficulties of quantifying the risks of such thinking, but it does not invalidate it.

While verification is intuitively appealing, it is not the most robust technique for arriving at the truth. In practice, evidence may be difficult or costly to gather. Karl Popper's (1959; 1974) theory of falsification and, indeed, Imre Lakatos' (1980) improvement of that theory, is an ideal of scientific progress that attempts to get round the pitfalls of verificationism. Yet, even there, there are problems. For example, the theory of gravitational waves, first proposed by Oliver Heaviside in 1893, and then by Henri Poincaré in 1905, and then predicted by Albert Einstein in 1916 in his general theory of relativity, had to wait until 1974 for the first indirect evidence, and until 2015 for the first direct observation, thanks to the Laser Interferometer Gravitational-Wave Observatory (LIGO) (LIGO Scientific Collaboration and Virgo Collaboration 2016). This demonstrates the pervasiveness of the problem of weighing two propositions, one with little or perhaps even no evidence, and another with an abundance of evidence.

Probabilistic reasoning is essential to the development of science. Although a quantitative view of probability is often what we think of when we ponder the importance of probabilistic reasoning to the hard sciences, in fact, some of the most important questions escape quantifiability. For our purposes, the class of questions that can only be qualitative rather than quantitative, are those that relate to weighing the probability of two theories. For instance, the big bang theory is believed to be more probable than steady-state theory, but the degree of plausibility is unknown (Franklin 2015).

The Duhem-Quine Thesis

Adding to the difficulties of weighing arguments is that, the more complex a theory, the harder it is to determine its relationship with an isolated piece of data. Scientists are forced to use the probability of ordinary language and logical tools to advance their arguments. It is for this reason that Pierre Duhem (1998, 260) stated that in physics, “an experiment …can never condemn an isolated hypothesis but a whole experimental group”. Willard Van Orman Quine (1951) extended the case for underdetermination in science. Quine argued that all human knowledge was a “man-made fabric” that only impinged on experience along the edges. This is tantamount to saying that human knowledge is necessarily only a subset of experience.

Human knowledge is an eternal quest to understand experience, or reality, and epistemologically, there is no difference, Quine says, from the physical objects science studies, to the gods of Homer. In fact, physical objects and gods of Homer are posited by us in our quest to understand human experience and guide us through the future. Therefore, all human knowledge, including logic and mathematics, can be revised in light of future experience[15].

The Duhem-Quine thesis, then, holds that in science, a hypothesis cannot be experimentally tested in isolation, because such a test makes various background assumptions, and this makes unambiguous Popperian-style falsification impossible. Indeed, no single hypothesis can make a prediction, instead, it does so as part of a bundle of hypotheses (hypothesis and background assumptions). We can only falsify the bundle of hypotheses and not the hypothesis alone. So, if we believe that the background assumptions are true, we will have

[15] Quine would later revise his position and exclude logic from this, saying that it would be “changing the subject” if logic was revised due to experience. Nevertheless, his larger point remains.

bounded reasons for believing that the hypothesis being tested is wrong if the scientific prediction fails.

Monotonicity

The pre-Socratics recognised that rationally persuasive arguments could be divided into the necessary, probable, and the persuasive yet sophistical, and Aristotle built his non-deductive logic upon that distinction. Monotonicity is a fundamental property of classic logic. Under monotonicity, a hypothesis of derived fact is extensible as more assumptions are added. "All men are mortal. Socrates is a man. Therefore Socrates is mortal", is perhaps the most famous example of this. Using the inference rule, weakening, this can become, “All men are mortal. Socrates is a man. Spiders weave webs. Therefore Socrates is mortal”. The addition of new premises does not alter the conclusion. Monotonic logic does not allow for reasoning by default in which there are default assumptions; abductive reasoning, in which the most likely explanation is sought; and belief revision. Non-monotonic logic, on the other hand, allows people as well as artificial intelligence, to conduct defeasible inference[16] and draw a conclusion about a situation or problem despite having incomplete information (Strasser and Antonelli 2019; Donini et al. 1990; Franklin 2015). This allows them to change their conclusion with new information. Bayesian[17] statistics is an example of this, but not every instance of non-monotonic reasoning is numerical. Jewish and Roman law explicitly treated non-monotonic logic and was present in the mediaeval theory of presumptions. In artificial intelligence, there is some argument as to whether the weights assigned to arguments should be quantified.

Scaling Arguments

[16] This is rationally persuasive but not deductively valid reasoning.

[17] The Bayesian approach was pioneered by Thomas Bayes (1763) and extended by Pierre-Simon Laplace (Stigler 1986). Rather than working with frequencies, one works with reasonable expectations. Bayesian statistics is arguably a part of propositional logic (Howson 2001), in which one reasons with propositions whose truth value is unknown. A hypothesis is given a probability weighting, whereas with frequency probability, it is not. Before testing, the probability assigned is known as the prior probability, and subsequent updates are posterior probabilities. Bayes did not provide a statement of his theorem, that was left to Laplace, who gave us,

$$P(B_i|A) = P(A|B_i)P(B_i)/\Sigma_j P(A|B_j)PB_j)$$

We also see that probabilistic arguments have and can be made by scaling arguments. In legal systems, concepts of proof beyond reasonable doubt, and moral certainty, have been advanced and established. Concepts of half-proofs exist, and arguments that are “more probable than its opposite”, and “almost certain”, have existed in the Greek, Roman and Talmudic tradition (Franklin 2015).

Statistical syllogisms

Statistical syllogisms are non-deductive syllogisms that employ inductive reasoning to come to a conclusion about a particular case from a major premise, also known as a generalisation, that is true for the most part. In doing so, one often uses qualifiers such as “most”, “frequently”, and “rarely”, or a statistical generalisation. When we travel by plane, we do so because we believe that “almost all” flights arrive safely. Confidence intervals are defended using statistical syllogisms (Franklin 1994). An example of statistical syllogism follows:

1. Most Popes reside in the Vatican
2. Francis is a Pope.
3. Therefore Francis resides in the Vatican.

The conclusion is logically supported by the premise, and yet the conclusion may be wrong, even if that is unlikely. Statistical syllogisms can be generalised as being of the form,

1. X proposition of α and β
2. I is an α
3. I is a β

where α is the reference class, β the attribute class and I the individual object.

Non-numerical statistical syllogisms of the “most part” kind exist in Aristotle (350 AD), where he refers to "that which people know to happen or not to happen, or to be or not to be, mostly in a particular way, is likely, for example, that the envious are malevolent or that those who are loved are affectionate." The Talmud urges people to “follow the majority” when in doubt (Franklin 2015). Not only did ancient medicine make use of statistical syllogisms, but even today, drugs are designed to cure a certain proportion of cases, so that the drug can be offered to all sufferers of the disease. Statistical syllogisms may even be used in law, although a judgement cannot solely be based on them because of the undue burden they place upon a defendant who finds himself within a reference class. For example, if there is a riot and 100 are observed in a city square, and it is known that 70 of those people rioted, smashed shop windows and stole goods, the following statistical syllogism can be formed,

1. 70 of the 100 people in the city square rioted
2. The defendant was in the city square
3. Therefore, the defendant is likely to have been one of the rioters

In the ancient tradition, statistical syllogisms were largely non-numerical. In the late Scholastic tradition, the relationship between proportions of returning ships or dead annuitants and insurance rates and annuities, was understood (Franklin 2015).

Combining Arguments

The ancients were also used to combining evidence to strengthen an argument. They understood that even though each evidentiary item was weak, the combination of items could lead to a strong argument. This was understood in physiognomics and law, and by Hipparchus and succeeding astronomers (Franklin 2015). The development of the method of least squares is an example of how scientists learnt to get closer to truth in the face of errors.

Statistical Evidence

The ancients also understood the importance of statistical evidence, although they did not attempt to quantify this notion. For instance, Aristotle, Cicero and the Talmud all understood that a long chain of evidence made chance as an explanation less and less persuasive (Franklin 2015). This is understood as well in modern legal systems where a defendant's prior criminal record, for example, lends credibility to the charges at hand.

Frank Knight's Theory of Risk and Uncertainty

A Theory of Knowledge

Knight (1964) developed his notion of risk and uncertainty as part of his pure theory of profit, arguing that it was under uncertainty that profits are made. He takes as axiomatic the notion that human action projects itself into the future, therefore, human existence depends upon extracting some knowledge of the future. The problems of life arise from the limits of that knowledge, because, while we do not act from pure ignorance, we also do not act from perfect knowledge, rather, we act from partial knowledge about the future. The future is uncertain, but it is not wholly impenetrable.

Subjectivity of Knowledge

Without using the word, Knight argues that our partial knowledge of the future is *subjective*, it is based on our "opinions" and perceptions. Yet, despite acknowledging this subjectivity, Knight (1964, 198-199) says the following,

> "...it is unnecessary to perfect, profitless imputation that particular occurrences be foreseeable, if only all the alternative possibilities are known and the probability of the occurrence of each can be accurately ascertained. Even though the business man could not know in advance the results of individual ventures, he could operate and base his competitive offers upon accurate foreknowledge of the future if quantitative knowledge of the probability of every possible outcome can be had."

It is rather hard to square the subjectivity of our beliefs about the future, with the assumption that it is possible to know "all the alternative possibilities" and to be able to assess the probability of each, an assumption that is only possible if one has complete, rather than partial knowledge. Knight's hypothetical assumes a degree of prescience that is surely beyond the ordinary businessman, as he claims that "accurate foreknowledge" is possible given a "quantitative knowledge" of probability distributions. Outside of highly controlled conditions with tractable randomness such as the games out of which expected value theory was built, it is absurd to speak with Knight's degree of confidence about "measurability" and even more absurd to speak of a computable probability seemingly free of an error rate. Outside of those highly controlled spheres, probabilities are hard to estimate, not simply because of the inherent unknowableness of the future, but also because of errors in the data. Yet, Knight's distinction between risk and uncertainty seems to forswear the possibility of an error rate. No earthly person can, then, be said to deal with Knightian risk. Human action implies difficulties in assessing probabilities and therefore the constant presence of errors in estimation. Thus, there is uncertainty about one's risks, about what alternative paths one faces, and what the probability of each is.

This can be demonstrated by extending Kurt Gödel's (1992) incompleteness theorem[18], in a very naive way, it must be said, with this thought-experiment: suppose we had the ability to amass all the raw data in the world and to optimise and analyse it in the most advanced computer imaginable, to form investable theories. We would, however, hit a roadblock: investment theories, being the subject of the program, could not be part of the structure of the language of the program. We would have to make unverifiable and possibly untrue assertions about investing. So long as there are thinking people, raw, or empirical facts disappear and what is left are value-determined facts. Values encompass standards of evidence and rules of inference, and ensure that even with a perfect and complete dataset, there will always be uncertainty arising from our own thoughts. Thus, it is impossible to experience Knightian risk, because risk itself is uncertain[19].

On Error in the Estimation of Risk

Knight acknowledges the difficulties of inferring the impact of our actions on the future, and of inferring what the future would look like without our actions, characterising this process of inference as neither being infallible, or "even accurate and complete"[20]. This is of course against the more confident tenor of his more famous statement, which bowdlerised uncertainty from risk. Indeed, Knight acknowledges that we do not even understand the present in its fullness. While Knight's view of risk has reduced it to an error-free computational problem, wherein it is possible to achieve an accurate estimation of the future, his discussion of risk finds that it is indeed strewn with uncertainty, saying,

> "We do not perceive the present as it is and in its totality, nor do we infer the future from the present with any high degree of dependability, nor yet do we accurately know the consequences of our own actions. In addition, there is a fourth source of error to be taken into account, for we do not execute actions in the precise form in which they are imaged and willed." (Knight 1964, 202).

---

[18] Every arithmetic system contains propositions that can neither be proven nor disproved.

[19] This suggests a typology of risks: black swans, discussed by Taleb (2010); grey rhinos (highly probable, high impact yet neglected threats) discussed by Michelle Wucker (2016); white swans, which Nouriel Roubini and Stephen Mihm (2011) argue define most financial crises, that is, they are high impact, yet predictable results of structural factors; or dragon kinds, proposed by Didier Sornette (2009), and which are high impact, somewhat predictable events.

[20] Knight 1964, 202

The Knightian view of risk and uncertainty that exists today is a much less sophisticated version of what Knight had in mind. While it may be possible to quantify an error rate, that quantification would be normative, subjective, and possibly very wrong, and therefore, it would not remove the uncertainty surrounding risk. In addition, Knight's discussion of the role of consciousness shows the fallibility of decision making and risk assessments. For Knight, consciousness exists to make inferences and those inferences are, he assumes, rational. That rationality does not imply the absence of error. Errors are made because information is both incomplete and complex.

The present and future are linked together through the principle of causality. Reasoning, Knight says, rests upon the principle of analogy. We reason by assuming that moments in time are analogous, that constant modes of behaviour between things will continue into the future, and that the future will look at least in some part like the past, because we believe that all these things are causally related. From this, Knight concludes that, *ceteris paribus*, "the same things always behave in the same way" (Knight 1964, 205). It is the introduction of change that disrupts the easy relation between things.

The demands of inference and forecasting force us to confront the problem of classifying things, and of figuring out the conditions that govern those things. From there, it is possible to forecast future behaviour. These are all, of course, not questions of mere computation. Difficulty arises from the concept of a "thing": Knight notes that these conditions, or "circumstances" to use his word, are often made up of "other things and their behaviour" (Knight 1964, 205). How much of the universe, Knights asks, is made up of things whose character never changes? Other writers have treated that question without directly referring to Knight: when Nicholas Nassim Taleb (2010) speaks about Black Swans, he asserts that Black Swans are more frequent than we believe and are the most consequential risk events we experience; whereas, to take another example, when Michelle Wucker (2016) discusses grey rhinos, she asserts that predictable events are plentiful, have high impact, but are often ignored.

Analogy works because there is a finite number of possible classifications of things. If it were not so, there would be no use in classifying anything. The reality of classes is possible because associations are not randomly created, there is some regularity. Classification holds even if the things within that classification are different in many ways. We assume, because experience tells us so, that in general, the same things within each class will behave the same way under the same conditions. However, given the uniqueness of each thing, and the number of features each thing has, rather than belonging to single classes, things move from class to class depending on the relevant problem. For example, a person may be “French” in a problem of national identities, but, “chess player” in a problem of professions. Therefore, classes are governed by “modes of resemblance" (Knight 1964, 205).

The principle of analogy allows finite intelligences, that is, intelligences that cannot grasp the wealth and complexity of information in the universe, to make sense of the world. In ventilating the theory of knowledge thus, Knight is at pains to distance himself from the irrationalistic tendency of post-Bergsonian theories of knowledge. In the search for exact knowledge, this is the only valid means of coming close to it. Given causality, the problem of knowledge is not a problem of randomness, that is, of things randomly associating with each other, but a problem of the vastness of that knowledge and its complexity, and the challenges of dealing with that[21]. Yet, the power of analogy is such that we are able to act through the world with a quite staggering amount of ignorance about the things around us.

Although logic strains for exact knowledge, and universally valid scientific assertions, Knight tells us that we are largely forced to rely on imprecise, weak assertions of the form, “some X is Y”, which, although allergic to attaining the status of scientific laws, do have what Knight describes as “scientific utility”, when the numerical proposition of instances in which X is associated with Y, is known. The essential principle of insurance is grounded upon this kind of weak reasoning. Knight illustrates the power of this weak form of reasoning through statistical syllogisms, by pointing out that although a bottle maker is forced to accept some proportion of bottles bursting, and will of course not know which particular bottles will burst, this knowledge allows the bottle maker to convert the loss into a fixed cost. This form of reasoning of course only uses a numerical proportion where the situation is important enough to justify assuming the costs and difficulties of collecting sufficient information, and determining the strength of the association. In other situations, one makes do with a vague, non-numeric sense of proportionality.

Knight on Probability Situations

While Knight is famous for his delineation of risk and uncertainty, his theory of risk and uncertainty is better seen as a theory of “probability situations”, in which we use a

[21] One could say that “randomness” is an epistemic category rather than inherent to the nature of things. In other words, when we say something happened “randomly”, it is not because that event was truly random, that we call it “random”, but because we are unaware of the causes behind that event.

combination of *a priori* probability, statistical probability, and estimates to navigate through risk and uncertainty. *A priori* and statistical probability exist to determine the numerical proportions by which X is associated with Y. Estimates are made to bridge the divide between ignorance and knowledge under uncertainty.

A priori probability is the kind of probability we know from games of chance and in logical and mathematical theories of probability. The homogeneity of the instances or events, which exists except for "really indeterminate factors", gives *a priori* probabilities the force of mathematical propositions. It has the advantage of arriving at a notion of the objective, or "real probability" of something in advance of the event and without calculation. For example, we know the real probability of getting a one when we throw a die, without having to calculate this, because of the physical properties of the die.

On the other hand, the risk of a building burning cannot be ascertained *a priori*, it must be calculated. In business, Knight says, we never meet with *a priori* probability, but we do meet with statistical probability. In business, we cannot know in advance what the distribution of events will be, however, we can, to some extent, use past experience as a guide to what this distribution *may* be. Statistical probability makes generalisations about a reference class. This statistical approach can, in some instances, lead to a good deal of certainty and in others, that is not possible. The nature of this statistical approach means that it cannot give a definitive answer. We do however make an *a priori* assumption that past relationships will hold to some degree in the future. Furthermore, business conditions are not, like games of chance, homogenous, and thus subjective judgement must be applied to data, in, for example, putting an object within the appropriate reference class. Nevertheless, each event is unique, for example, no two fires are precisely alike, we can try and achieve a good deal of homogeneity within a reference class, by creating very granular reference classes, yet, even there, it is not possible to achieve a notion of real probability. Statistical probability carries, as Knight notes, a logical problem: to achieve absolute homogeneity, it must have absolute uniformity, and thus, probability becomes impossible. In other words, the most granular reference class is one which only has a single case, given the uniqueness of cases, and so, we cannot speak about "some X is Y". Statistical probability presents one with two problems: it is not possible to remove all somewhat determinate causes; and, the impossibility of articulating all equiprobable alternatives.

A priori and statistical probability are related to risk situations. Knight posits a third probability, "estimates", which governs situations of uncertainty. Broadly, for Knight, estimates are subjective, one does not have a "valid basis" for one's classification of events.

One may very well assign a number to one's judgement of the likelihood of something, but that number has limited objective content, consequently, one is likely to err. When economists and investors speak about "Knightian uncertainty", they are referring to the third of his trichotomy of probability situations.

Indeterminateness and Knowability

In discussing estimates, he gives the earliest known treatment of the problem of red and black balls in an urn. The probability of red and black balls depends upon the information that one has: a person who only knows that there are red and black balls might posit that there is an even split of the two. A person who knows that the ratio is 3:1 may conclude that there is a seventy-five to twenty-five split. The "real probability" may be the second, but in the face of uncertainty, it would be rational to suppose that it was actually the first. The problem of the urn is a problem where indeterminateness is due to our ignorance: the real probability is knowable.

In fact, Knight points out that where knowledge is complete, there is no probability, only certainty. Real probability, Knight believed, depends upon the inherent unknowability of causes, rather than our ignorance of them. That said, we cannot assume that events will follow the law of indifference, and so, we are obliged to build an empirical foundation for our probability. Of course, Knight realises the tension between the fact that probability derives from the properties of things, and therefore, determined, and yet, we reason as if the probabilities are indeterminate, with unknowable causes that follow the principle of indifference.

The mathematical theory of probability assumes, Knight says, an unbridgeable divide between complete determination and complete indifference, and so, elementary probabilities are always equal. More simply, it assumes that, given indeterminacy, or uncertainty, we must start off our probabilistic inquiries by assuming that outcomes are equally likely.

Throughout his discussion, Knight takes the view that probabilities have no cause, saying, for instance, that,

> We do not merely feel that we know no reason why the coin shall fall heads or tails; we know in a positive sense that there is no reason, and only under this condition do we make the probability judgment with any confidence. (Knight 1964, 222)

This is a very strange position to take. Causality is so axiomatic that it is embedded in the very structure of our language (Copley and Martin 2015; Mackie 1980). It seems odd to dismiss it so quickly and with so little argument, especially given that, in developing his ideas on knowledge and consciousness, he holds causality as axiomatic. One suspects that this Knight's framing of the issue is merely clumsy, for there are obviously "reasons" why heads or tails fall. Knight confuses causality with determinacy. Max Born (1949) showed that determination should be separated from causality, explaining that determination relates to our ability to make robust forecasts and retroductions based on available data, because of the way things are linked together by the laws of nature, whereas causation can be probabilistic, or "nomic", or singular. It is not necessary to abandon causality in order to promote indeterminacy.

The Limits of Knight's Theory of Probability Situations

A priori probability, statistical probability and estimates, seem to exist along a continuum of homogeneity, from situations of absolute homogeneity, governed by *a priori* probability, to situations of purely unique instances, governed by estimates (Runde 1998). In arguing his case for the presence of statistical probability and estimates in business, he shows that error is a feature of operations, because of our inability to derive the "real probability" of any event. These probability situations are, therefore, also continuums of error, with our error rate increasing with the diminishing of homogeneity. However, his theory is somewhat self-contradictory, because, whereas he defines these probability situations along a continuum of homogeneity, he also defines them as being *logically* different, split by *a priori* and *a posteriori* methods (Runde 1998). The problems with his definition mean that we cannot properly map his trichotomy of probability situations, onto his dichotomy of types of chance. Jochen Runde (1998) argues rather convincingly that the perfect homogeneity of instances, or trials, is not a sufficient condition for being able to calculate *a priori* probabilities, even if all possible alternatives, or outcomes, are known. For example, the probability of outcomes of a throw of a weighted die, cannot be determined *a priori*: it is only when outcomes are equiprobable that they can be determined *a priori*. Runde locates Knight's difficulty in his confusion of instance and alternatives, or, in modern parlance, trials and outcomes, which led to his conclusion that the homogeneity of trials was responsible for the difference in probability situations. Runde proposes that we redefine Knight's probability situations thus:

1. *A priori* probability should refer to situations where probabilities can be calculated from general principles, and are, thus, equiprobable and mutually exclusive.
2. Statistical probability should refer to those probabilities calculated from empirical observation of the frequency of outcomes in largely homogenous trials.
3. Estimates should refer to non-quantifiable probabilities, either because of an insufficiency of homogenous trials, or because the probabilities cannot be inferred from general principles.

In this way, the problems of Knight's trichotomy are fixed, his tendency to associate mathematical probability with equiprobable outcomes is gotten rid of, along with his assumption that statistical probability is so because of the lack of homogeneity of trials.

Von Mises' Theory of Probability

Von Mises' Scepticism About the Use of Mathematics

Although he built upon Knight's theory of risk and uncertainty, accepting the distinction between risk and uncertainty and the origin of profits in uncertainty, Ludwig von Mises (1998) was less enamoured with the use of mathematics in these realms. Indeed, he remarked,

> "When the Chevalier de Méré consulted Pascal on the problems involved in the games of dice, the great mathematician should have frankly told his friend the truth, namely, that mathematics cannot be of any use to the gambler in a game of pure chance. … People suspected that the puzzling formulas contain some important revelations, hidden to the uninitiated; they got the impression that a scientific method of gambling exists and that the esoteric teachings of mathematics provide a key for winning." (von Mises 1998, 106)

Von Mises believed that mathematicians had confused the field of probability, and that the conflation of mathematical probability with probability as dangerous and limiting, saying,

> The problem of probable inference is much bigger than those problems which constitute the field of the calculus of probability. Only preoccupation with the mathematical treatment could result in the prejudice that probability always means frequency. (Von Mises, 1998, 107).

Furthermore, von Mises saw a further conflation, between probabilistic thinking and inductive thinking of the natural sciences, as problematic. This seems to relate to his belief that as an *a priori* domain, it was erroneous to attempt to derive economic laws from empirical observations[22]. He opposed the prejudice that equated probability with frequency and the confusion of probability with induction[23], the great *bête noire* of many philosophers, with David Hume (1739; 1748), and Karl Popper (1974), the most famous antagonists of the use of induction.

Von Mises' saw the essence of insurance as the "pooling and distribution of risks". Probability and the knowledge of probability distributions, is not an inherent feature of insurance. This is similar to the views that Knight held about the pure form of insurance. Around the time that von Mises was writing, the importance of probability in insurance was hotly debated, with thinkers such as Baptist, de Finetti, and Berger feeling that probability plays an important role, while thinkers such as Hammon, Clarke, Lah doubted its important and others such as Hagstroe, Shannon and ten Pas rejected its importance entirely (C.C. Nicholl 1946). From a practical point of view, the difference in profitability between two insurers hinges in part on their ability to gain probabilistic insights from data. His broader point that we naturally insure goods without recourse to probability and that conceptually, probability is not implicated in insurance, is undeniably true, but that does not speak to what insurance is at heart.

Time and Risk

He proposed that risk and uncertainty exist because man works through time. Without time, there is no risk, no uncertainty, simply finality. Human action through time is inherently uncertain and shapes the nature of risk (Bernstein 1996; von Mises 1998). Time is even more important when the decisions we make cannot be reversed, or can only be reversed at great difficulty (Bernstein 1996).

[22] This can be seen through the fact-value dichotomy. For instance, one cannot answer the question, "What is rain?", with empirical observations. One needs notions of the category "rain", in order to know what the empirical observations mean.

[23] Von Mises seems to refer to inductive reasoning, rather than mathematical induction. Inductive reasoning being the process by which a series of observations leads to the establishment of some rule or principle, and mathematical induction being, despite its name, a rigorous deductive process (Jaynes 2003).

Even if the future is knowable to some higher intelligence, it remains a mystery to humanity: the human actor does not have access to the knowledge that would make the strands of the future perfectly clear. His is a realm absent of apodictic certainty, regardless of whether the future is predetermined or not. In other words, our probability distributions, even if known, always carry an error rate, and are, thus, uncertain. If it was possible to know the size of the error rate in advance, it would not be a true error rate.

Defining Praxeology

Von Mises distinguishes the natural sciences from what he calls “praxeology”[24], and it is around this distinction that he argues about the limits of the knowability of the future. By “praxeology” is meant the theory of human action, which assumes the intentionality or purposefulness of human action. Economics is a praxeological subject that became possible as mankind discovered regularities in the market, and this necessitates a distinct methodology to that of the natural sciences. Economic activity is about deciding between a range of choices. As such, von Mises rejected attempts to make economics a science, as much as the Austrian School had previously rejected the historicism of the Younger-branch of the German Historical School during the Methodenstreit[25]. Praxeology assumes that thinking is not just purposeful, but also logical. If thinking is illogical, it is not thinking[26]. As a praxeological subject, von Mises held that economics was, therefore, an *a priori* subject as well, whose laws were deductively derivable. Von Mises characterised “the problem of truth and certainty” as being epistemic and the problem of probability as being praxeological.

Reflexivity, Indeterminacy, and Uncertainty

[24] The term derives from Alfred Espinas and the Austrian and Polish schools, led by von Mises and Tadeusz Kotarbiński respectively, developed theories of human action (Ryan, Nahser, and Gasparski 2002; von Mises 1998).

[25] German for “method dispute”. It was a period in the 1880s when the Austrian School and the Younger-branch of the German Historical School debated the role of general theory in the social sciences, and the importance of history in understanding human action. It also involved debates around the state and the individual, although it was largely centred on methodological questions, with Carl Menger refining the Austrian approach and Gustav von Schmoller the Historical School’s approach.

[26] Roderick T. Long (2004) offers a powerful defence of this approach and Gerd Gigerenzer (2018) provides a strong argument that the plethora of biases in behavioural economics are a sign of problems with the field rather than with human thought.

The use of probabilistic reasoning with the soft sciences and history presents mathematicians with a number of problems, as von Mises rightly realised. Soros (2013; 1987) believed that the distinction between the hard and social sciences was due to the presence of reflexive processes within social systems. By reflexive processes are meant those processes that explain situations in which there are thinking participants, and where those participants shape reality. Soros advanced two propositions:

1. The principle of fallibility: the perceptions of thinking participants do not precisely correspond with reality. Perceptions, especially complex ideas, are always partial[27], or inconsistent or both.
2. The principle of reflexivity: the imperfect views of thinking participants, exercised through a "cognitive function", shape the situation that is being analysed, through a "manipulative function". This connects subjective reality with objective reality. Working together, these two functions can interfere with each other, as the independent variable of one is the dependent variable of the other, creating a recursive relationship. The self-reference of social systems creates indeterminacy and uncertainty.

This creates a myriad of problems for situations in which there are thinking participants. It is impossible to use quantitative probabilistic reasoning in the same way that it is used in the hard sciences, because thinking affects reality in ways that defy probabilistic reasoning. This is why it is so hard to predict the future. For example, meme investors who bet on stocks such as AMC, and Hertz, invested in companies that were fundamentally "broken". Yet, that investment allowed AMC to escape bankruptcy, and helped Hertz to rapidly emerge from bankruptcy. This is a refinement of the value investing notion that intrinsic firm value determines long-term market value (Graham and Dodd 2009; Graham 2006). It is a world without fixed reference points or independent variables.

Although Soros felt that reflexive systems separated social sciences from hard sciences, Eric D. Beinhocker (2014) proposed that reflexive processes should be seen within the literature of complex adaptive systems, and that systems exist along a "spectrum of complexity" and that under certain circumstances, natural and artificial systems can be "complex reflexive" like social systems. So, rather than seeing a fundamental break between

[27] "Bias" is the more used word, but in this instance, bias would suggest Kahneman and Tversky (1979) style bias, whereas in this instance, one means merely that human thinking is incapable of encompassing all of reality.

natural sciences and social sciences, it is truer to say that the real differences are in terms of where a discipline lies on that continuum.

An adaptive system is one in which there are thoughtful, interacting agents. It becomes a complex adaptive system when its behaviour becomes difficult to model due to the nature of the interactions between the adaptive system's components or between the adaptive system and its environment (Miller and Page 2007; Beinhocker 2014). Complex systems are characterised by features such as their spontaneous order, nonlinearity, feedback loops, emergence and adaptation. Importantly, they occur in a number of domains and are not restricted to the praxeological fields that von Mises posits. For instance, complex systems have been studied as part of chaos theory in deterministic systems, nonlinear dynamic systems, neural networks, dissipative structures, and quantum chemistry. Adam Smith's notion of the "invisible hand" reflects the spontaneous order that marks complex systems. F. A. Hayek (2014) noted that market order was "the result of human action but not of human design". Von Mises' (2012) own insight into the economic calculation problem is an example of the computational complexity of social systems. Nevertheless, it was not until the 1970s that complex systems were studied as a separate subject, and it was not until the founding of the Santa Fe Institute in 1984, that there was a research body dedicated to studying complex systems, with a multidisciplinary set of participants such as Kenneth Arrow, Murray Gell-Mann, and Phil Anderson. Since that time, there has been more multidisciplinary collaboration between mathematical physicists and economists, and mathematical physicists have begun to study economic phenomena[28].

Cross et al. (2014) also argue that Soros' model doesn't completely segregate the social and the hard sciences. Furthermore, they argue that participants would be subject to the Duhem-Quine entanglement of hypothesis and background assumptions, as they try to falsify their hypotheses. Soros' theory opens the door to far-from-equilibrium economics, that comes into play when equilibration is not instantaneous, and where mispricing, when corrected, does not result in a change in the economic structure built upon that mispricing.

On Case and Class Probabilities

Von Mises characterises a probabilistic statement as one for which we do not have complete knowledge about to make a firm determination as to whether it is true or not. Of these

[28] Peters and Gell-Mann (2016) is an example of one such intervention into economics by mathematical physicists.

probabilistic statements, he develops two categories: class probabilities, which relate to frequencies, and case probabilities, which is the probability of unique events that cannot be placed into reference classes. He believed that natural sciences are dominated by class probabilities and the praxeological sciences by case probabilities.

With class probabilities, we know or think we know everything about the reference class[29], but of individual members, we know nothing beyond their membership in the reference class. With case probabilities one knows some of the characteristics of the individual event, but this island of knowledge exists within a sea of incomplete knowledge. Case probabilities are the realm of a kind of “Knightian uncertainty”. As with class probability, von Mises recognises the incompleteness of our knowledge. Von Mises’ view of the mathematical theory of probability led him to believe that it had no place in the study of case probabilities.

In dealing with the world of class probability, von Mises posits that we use the tools of understanding, which are “always based on incomplete knowledge” (von Mises, 1998, 112). He does not develop this idea further, but it seems that he is referring to the kind of ordinary language reasoning that we have discussed in earlier sections, although he does not link this with any notion of logical probability.

Von Mises gives a number of other examples of case probability. For instance, he argues that the result of an election is, properly, a unique event and that it is wrong to use a frequentist approach to predict the winner of a presidential election. Superficially, we can say that even an uncertain election has some level of built in certainty given the electoral rules and structural drivers of election results. Von Mises was a determinist and would agree that everything has a prior cause. For instance, in the United States’ two-party system, the possible winners are narrowed down to two people, one of the Republican and Democratic Party nominees, and we can predict with a high degree of certainty, all but the swing states. Furthermore, with a fairly high degree of confidence, once the party nominees are chosen, one already has near-certainty (death or a nominee stepping down aside) of who will be the president. The degree of uncertainty surrounding the question of who will be the president is

[29] A reference class problem exists prior to calculating the probability of a single case. One has to determine what the appropriate reference class is, and this carries the risk of error. Each individual object can be logically placed in a multitude of different reference classes. John Venn (1888) pointed out that “It is obvious that every individual thing or event has an indefinite number of properties or attributes observable in it, and might therefore be considered as belonging to an indefinite num- ber of different classes of things.” Therefore, there is a non-numerical element mathematical probability, which, as discussed in previous sections, one cannot assign a number.

very low once the nominees are clear, if we think of the problem as, “Which American will become president?”. However, as a question of which of the two main candidates will become president, the uncertainty is much sharper than first seems. Bruno De Finetti (1937) suggested that a subjective probability statement defines the price, known as the "operational subjective probability", at which the gambler would be willing to bet on that outcome. Such a gamble must not include the certainty of loss through arbitrage by one’s counter-party, otherwise the probability statement becomes incoherent[30]. Now, the more time to an election, the greater the uncertainty and opportunity for unpredictable events and even predictable-but-ignored threats to surface. For instance, in the weeks leading up to the 2016 election, Senator Hillary Clinton looked set to win, but she faced a “black swan event” in the form of the FBI Director, James Comey, making a statement about an investigation about her. Taleb and Madeka (2019) extend De Finetti’s result in their discussion of election forecasts evolving through time, pointing to the equivalence of election forecasts to pricing binary options. Counter-intuitively, the effect of the heightened uncertainty the further along one is from an election, is to give the chance of any candidate from winning the election just a 50% probability, although, theoretically, it can be anything, and, that forecast should not move around a lot until very shortly before the Election Day. This 50-50 weighting and low forecast volatility must hold, otherwise, the gambler will violate De Finetti’s coherence criterion for making such a forecast, because the higher volatility allows for inter-temporal arbitrage as one trades with the forecasting “assessor”. Von Mises is right about the role of uncertainty in making forecasts meaningless, but those forecasts can achieve some meaning and quantifiability and be more properly assessed if there is “skin in the game” on the part of a person making a forecast. No forecast, however, has any probative value in terms of what the future holds, because of the radical uncertainty wherein, it only tells us what we believe about the future. Von Mises’ views on forecasting reflect his understanding that probability statements reflect our understanding of the world, and not the world, which is to say, they are subjective.

Von Mises and the Frequentist Interpretation of Probability

Von Mises observes that a doctor who knows that the chances of recovery for some disease are 70% based on known recovery rates, is working with the frequency of recovery for a class of sufferers of that disease. If that patient were to ask if the number of deaths is full, he

[30] De Finetti’s argument is at the heart of the solution for the fundamental theorem of asset pricing, which tells us that when prices are consistent, they form a set of probabilities, such that the equations of probabilities are satisfied.

would, von Mises points out, be confusing class and case probability. Without saying so, von Mises (1998, 111) then gives an example of Bayesian updating[31], of reasonable expectations[32], observing that,

> If the doctor himself treats the patient, he may have a different opinion. The patient is a young, vigorous man; he was in good health before he was taken with the illness. In such cases, the doctor may think, the mortality figures are lower; the chances for this patient are not 7:3, but 9:1. The logical approach remains the same, although it may be based not on a collection of statistical data, but simply on a more or less exact resume of the doctor's own experience with previous cases.

The curiosity here is that von Mises seems to recognise some form of Bayesian statistics, while conflating the frequency interpretation of probability with the mathematical theory of probability. This may be due to the fact that Bayesian statistics[33] waned in popularity at the start of the twentieth century. In fact, Bayesian probability, or what was for most of its history known as the "inverse method", dominated statistical thinking until the 1920s, after the 1922 publication of Ronald Fisher's "On the Mathematical Foundations of Theoretical Statistics". Fisher developed the first systematic frequentist approach (Fienberg 2006). In fact, the term "Kollektivmasslehre", or "collective" was first used by Gustav Fechner as he tried to establish a systematic frequentist approach (Fienberg 2006). It may be that, because the frequentist approach had come to dominate statistics despite the work of Harold Jeffreys, Abraham Wald, and Leonard J. Savage, that, by the time von Mises wrote *Human Action* in 1940, it was easy to conflate the frequentist approach with mathematical probability. Since the 1980s, there have been great advances in Bayesian statistics, triggered by the emergence of Markov chain Monte Carlo methods that have made Bayesian statistics more tractable and fueled the rise of machine learning and other developments (Bishop 2006).

The genealogy of von Mises' views on probability can be traced to his brother, Richard von Mises, who in his book, *Probability, Statistics and Truth*, developed an explicitly frequentist theory of probability. Not only does he build upon his brother's work, he even uses the same

[31] Von Mises does not mention Bayesian statistics at all in *Human Action*.

[32] In truth, frequentism is only one strand of probability, the others being subjectivism, propensity theory and logical probability.

[33] Bayesian thinking is a mathematical approach to navigating situations of uncertainty in which one has to formulate and test hypotheses. Indeed, the proposition that there is an "unknown" is itself a hypothesis about which we have some degree of certainty. These hypotheses do not reflect the nature of reality, they are simply a comment about our evolving state of knowledge or personal beliefs.

arguments. For instance, his examples of a lottery and mortality tables mirror that of Richard von Mises (1981, 16-17, 19-20). Moreover, von Mises' definition of a class probability is similar to his brother's definition of a "collective" probability[34]. Finally, the brothers shared the belief that it was either impossible or misguided to assign a number to the probabilities of non-class/non-collective events. Ludwig von Mises departed from his brother in holding that non-class/non-collective events were part of probability theory, and that probability was linked to uncertainty, something his brother does not seem to have believed.

It can also be argued that the frequentist approach is a legitimate approach in certain praxeological events. For instance, football analysts regularly use measures such as expected goals (xG), and traders rely on measures of returns dispersion to tell them something about the riskiness of assets. There are many areas in which humans act where probabilities are assigned numbers.

Risk as an Epistemic Category

Von Mises' understanding suggests a nascent notion of complex adaptive systems and the challenges of predicting within them, given the constantly changing relations as each agent reacts to incoming information. Within this constantly changing environment, one cannot exercise the certainty that exists in less complex systems, and one will depend on one's understanding to guide one. That begs the question, "Where does uncertainty come from?"

Von Mises does not specifically define what his idea of probability is. However, some points can be made. Whereas his brother, Richard, was an indeterminist and believed that the universe's physical properties were the cause of uncertainty, following from Heisenberg, Ludwig von Mises saw uncertainty as the product of incomplete knowledge. In other words, a superior being would know the course of the future regardless of the universe's physical or ontological properties, whereas our limited understanding prevents us from clearing the fog of the future.

Axiomatically, the world is governed by causality, in other words, there are no uncaused events in the world, even if we are blind to those causes (Hoppe 2007). The frequentist interpretation assumes that uncertainty lies in the physical properties of the universe, but it also assumes causality. For instance, the current FiveThirtyEight forecast for the 2022

[34] Cf. Richard von Mises (1981, 11-12) for von Mises discussion of collective probabilities.

House midterm elections is that Republicans have a 69% chance of winning the House and Democrats a 31% chance[35]. This forecast assumes that certain factors defined in the model's methodology will likely cause Republicans to win the House. Indeed, the same is true of all probability interpretations: they assume causation as a means to forecast the future. We are forced, then, to assume a subjective view of probability, rather than a frequentist one. Probability can be seen as a subjective measure of our certainty about future events. Thus, there is no need to separate probability between case and class probability. Probability does not exist outside of us, it is an inherently subjective subject. De Finetti (2017, 3), with some flourish, posited that,

> "only subjective probabilities exist – that is, the degree of belief in the occurrence of an event attributed by a given person at a given instant and with a given set of information"

If we accept this view, then, we can indeed assign numbers to our probabilities, regardless of where we are on the range of complex adaptive systems, our certainty decreasing with the increase in the complexity of the system. Probability as an epistemic and not ontic creature is free to be assigned numbers. Given that von Mises adopted his brother's frequentist interpretation of probability, he was bound to reject its use in describing case probabilities.

Keynes Theory of Probability

John Maynard Keynes', whose views closely aligned with those of Knight, echoed Knight's distinction between risk and uncertainty in his 1937 defence of his *General Theory*, saying

> By "uncertain" knowledge, let me explain, I do not mean merely to distinguish what is known for certain from what is only probable. The game of roulette is not subject, in this sense, to uncertainty; nor is the prospect of a Victory bond being drawn. Or, again, the expectation of life is only slightly uncertain. Even the weather is only moderately uncertain. The sense in which I am using the term is that in which the prospect of a European war is uncertain, or the price of copper and the rate of interest twenty years hence, or the obsolescence of a new invention, or the position of private wealth- owners in the social system in 1970. About these matters there is

[35] https://projects.fivethirtyeight.com/2022-election-forecast/house/?cid=rrpromo

> no scientific basis on which to form any calculable probability whatsoever. We simply do not know. (Keynes 1937, 213-214)

However, it is to his earlier work, Keynes' *A Treatise on Probability*, from which we must adduce his views on risk and uncertainty. This is perhaps the most famous attempt by an economist to come up with a logical probability, and it inspired Rudolf Carnap (1962) and E.T. Jaynes' (2003). The treatise received the high praise of Bertrand Russell (2016), who called it, "undoubtedly the most important work on probability that has appeared for a very long time," opining that the "book as a whole is one which it is impossible to praise too highly."

Keynes and the Logical Foundations of Probability

Much like von Mises, Keynes does not give an explicit definition of probability, because he felt that probability, properly conceived, was subsumed in logic, and so, his work goes about establishing the logical foundations of probability theory. At best, he described "probability as comprising that part of logic which deals with arguments which are rational but not conclusive" (Keynes 2013, 241). This theory would be a sort of can opener, allowing us to see how intuitive knowledge could form the basis of rational belief about things for which we are largely uncertain.

The Weight of an Argument

Keynes' logical probability starts off from the same assumption that von Mises rightly made: that probability is epistemic and not ontic. It seems implicit that Keynes holds the same view as Carnap that all inductive reasoning is probabilistic reasoning. Probability statements, Keynes argued, are expressions of logical relationships, between a proposition p, and proposition h[36], a relationship that is much like that of logical consequence, although it is obviously weaker. Knowing h and the logical relationship between p and h, one is given to a certain degree of rational belief in p. Probability expresses the extent to which a hypothesis or conclusion is confirmed given some evidence or premises (Keynes 2013; Carnap 1962; Jaynes 2003). Our proposition p, is not in and of itself possessed of an objective probability: its probability shifts with the logical relationship it is in. Probability is not, therefore, simply a device to understand the frequencies of random variables, rather, it is also a set of rules of

[36] *h* is typically a conjunction of propositions.

plausible reasoning (Jaynes 2003). There are then probabilities which are quantifiable, and those which are unquantifiable, or qualitative and perhaps even unique (Keynes 2013; Carnap 1962; Jaynes 2003).

Russell (1948) notes that people ought to remember that when one says that there is a probability of *a*, P(*a*), what that actually means is that there is a probability of *a*, given information *h*, but given the difficulties this would force upon us, he recognises the power of *a*/*h*, the relation of a proposition *a*, to data *h*, that Keynes proposes as a measure of the weight of an argument, “the fundamental symbol of the book”, saying, “It is remarkable how often the logic and philosophy of mathematical concepts has gone astray through the employment of symbols which did not contain explicitly all the variables upon which their value depended”. That symbol is, indeed, the essence of Keynes’ argument, showing that the weight of an argument is determined by the knowledge one has, and that, given the objectivity of knowledge, everyone should be able to arrive at the same conclusion when presented with the same knowledge. When we know something for certain, $a/h = 1$, such that knowing *h* tells us what *a* is, giving us indirect knowledge of *a*, whereas under uncertainty, knowing *h* gives us just a partial belief in *a*, so that we cannot know a proposition without knowing its premise. Jacob Bernoulli (2005), Thomas Bayes (1763), and Pierre-Simon Laplace (1995) all shared the belief that probabilities were implicated in the formation of reasonable beliefs.

Keynes on Subjectivity and the Unquantifiability of Risks

Rivals in economics, Keynes and von Mises shared a belief that the majority of probabilistic statements cannot be assigned a number. Keynes went further, suggesting that given two unequal probabilities, it did not follow that one is greater than, less than, or equal to the other. Probabilities exist along a continuum between certainty ($a/h = 1$) and uncertainty ($a/h = 0$), but their order does not exist on a single series, but in parallel series. In one of these series, probability can indeed be measured, but elsewhere, the probabilities cannot be measured. He likens probabilities to similarities, giving the following analogy:

> “There are also, as in the case of probability, different orders of similarity. For instance, a book bound in blue morocco is more like a book bound in red morocco than if it were bound in blue calf; and a book bound in red calf is more like the book in red morocco than if it were in blue calf. But there may be no comparison between the

> degree of similarity which exists between books bound in red morocco and blue morocco, and that which exists between books bound in red morocco and red calf. This illustration deserves special attention, as the analogy between orders of similarity and probability is so great that its apprehension will greatly assist that of the ideas I wish to convey. We say that one argument is more probable than another (i.e. nearer to certainty) in the same kind of way as we can describe one object as more like than another to a standard object of comparison."

Measurable Probabilities and Words of Estimative Probability

So where can we measure probabilities? Keynes suggests that if there is a set of exclusive and exhaustive alternatives, the truth of a proposition and the truth of its contradictory, which are equiprobable, then we can measure probability. We can directly assess the equality of two probabilities if we accept the principle of indifference. The principle of indifference, or non-sufficient reason, is a principle for assigning epistemic probabilities where there is no relevant evidence. In such a scenario, a person must distribute their degree of belief evenly across all possible outcomes. Keynes gives the earliest version of what is now known as the Ellsberg paradox, after Daniel Ellsberg (1961), in which one is presented with an urn filled with black and white balls, the proportion of which is unknown. Given the principle of indifference, whereas it may seem intuitive to assume, *a priori*, that every possible ratio of white and black is equally likely, the truth is that each ball has an equal chance of being white or black.

Keynes' belief that other probabilities cannot be measured is subject to the same criticism as von Mises' similar notion: the subjectivity of beliefs does not prevent us from assigning a number to a belief. The distinction between beliefs held about subjects in which probabilities are empirically observable, and those where they are not, is not a barrier to enumerating the degree of belief, but it is a restraint on the certainty of the assigned number. It may be argued that assigning a number to a degree of belief is meaningless under the most extreme uncertainty, but even there, and perhaps in some ways, more so there, assigning a number is useful. Sherman Kent (1964), writing in a report for the Central Intelligence Agency, noted that there is uncertainty around the words that people use when estimating the probability of events. People interpret words such as "probable", "likely", and "almost certain" differently. This can lead to one person believing that they are communicating that an event has a probability of 60%, whereas the person they are communicating to believes they mean the probability is 80%, for example. Kent proposed that numbers be assigned to a defined list of "words of estimative probability" with corresponding probability ranges, which would more clearly communicate an intelligence agent's degree of belief, than words can. By assigning subjective quantifications to probabilities, ambiguities are removed. Scott Barclay (1977), in an assessment of how NATO officers interpreted words of estimative probability that did not have clearly defined numbers assigned to them, and found that they interpreted those words very differently. Below is reproduced the chart showing the interpretative range of those words of estimative probability:

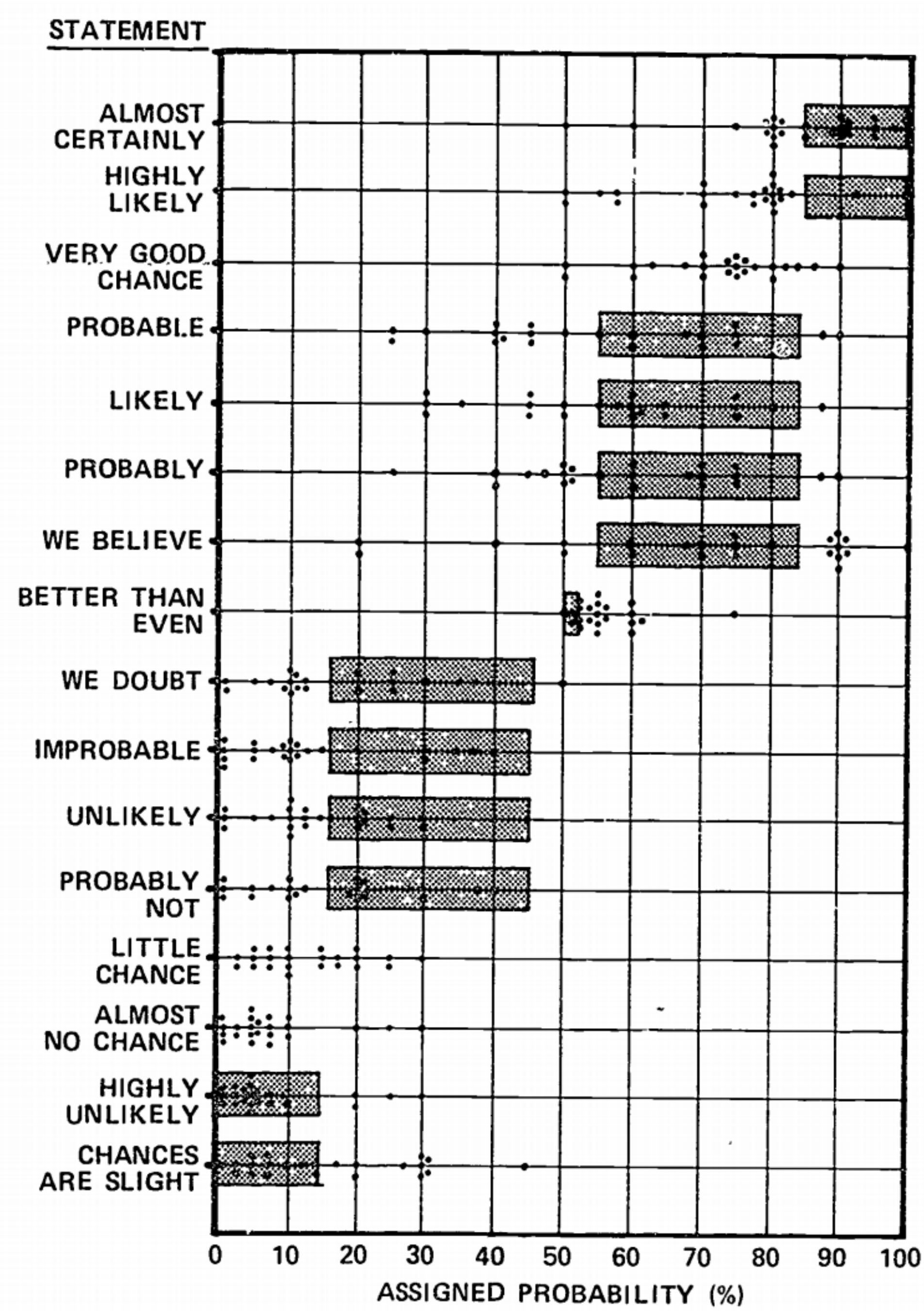

Figure 1 : The Meaning of Words of Estimative Probability. Source: (Barclay 1977, 68)

So, for instance, a report that uses the qualifier, “we believe”, led the NATO officers to assign an implied probability of anywhere between 55% and 85%. There is a clear overlap between “we doubt” and “probably”, and other inconsistencies abound[37]. This obviously has real world consequences in decision making. The confusion between the assignment of a number and the objectivity of that number, which led to a certain tendency to shun numbers, adds a layer of uncertainty to communications about uncertain events.

Keynes’ attempt to build a wholly subjective theory of probability is opposed by the existence of empirical measures of probability. There are categories of things whose probability is an objective part of their being, such as a fair die, or coin. It is clear that a frequentist approach has a place in many parts of science. If you stand beneath a falling rock, the chance is certain and objective that it will strike you. One can make the argument that even here, one can bring empirical measures of probability under the “subjective” umbrella by realising that empirical measures are not wholly objective: outside of thought, there are only raw facts, but so long as we are thinking about those raw facts, those raw facts can only be conceptualised by implicating our values, such as testing parameters, conceptual categories, and the like. Earlier, we extended Godel’s theorem saying that if we designed a program into which we would feed all the facts of the economy, in order to produce probabilistic statements about it, we would find that in order to design this program, we would have to explicitly and implicitly define the parameters, assumptions and categories of the program. So long as we are thinking, our subjectivity imposes itself even on the most objective of facts. The old question, “If a tree falls in a forest and no one is around to hear it”, tells us that without observation is nothing more than conjecture. However, in this case, we would be stretching the definition of “subjective” to its extreme limits.

Nevertheless, the nature of complex adaptive systems is such that we cannot create the idealised scientific conditions that make the frequentist approach so powerful in the laboratory. We must then depart from Keynes and take the Jaynesian position that the frequentist approach holds under certain, less complex conditions, and that Bayesian methods hold under complexity, where some kind of subjective procedure must be undertaken to weigh hypotheses. Thus while we might take the view that probability is a mathematical model that seeks to grasp some aspect of the world, the interpretation of probability used depends on where the subject sits on the continuum of complex adaptive systems.

[37] Crowd sourced platform, Probability Survey, provides a version of the survey provided to the NATO officers.

The Logic of Partial Belief

Whereas Keynes begins his analysis with the perception of a probability relationship, and ends it with a degree of belief, probabilistic logicians today, influenced by Frank P. Ramsey (1926), tend to begin with the degree of partial belief and ask themselves what probability relationship would justify such a belief. The difference stems from Keynes' argument that the probability relationship was objective, because knowledge is disembodied, outside of us. Like Keynes, Ramsey believed that probability is properly a part of logic, calling it, "the logic of partial belief and inconclusive judgement". This, Ramsey says, is not as important a revelation as that this notion of probability is not necessarily valid in statistical and physical science. Ramsey suggests that the rupture between frequentists and those who argue for the logical interpretation of probability is largely because they are speaking about different kinds of probability and the applications thereof. Ramsey criticised Keynes' idea that we should start from a probable inference and claim its objective validity, a notion which assumes that knowledge is disembodied, outside of us. Practically, what this implies is that if we have a certain degree of belief in the premise of an argument, we are led to one and only one degree of belief in the conclusion of that argument. Rather, Ramsey says, knowledge exists as part of a lattice of prior knowledge, and our beliefs are shaped by the knowledge we have and this is what leads to the degree of belief we have. This naturally leads to a subjective view of probability. Subjective probabilities are observed in the actions they cause under suitable circumstances[38]. Inferring the degree of belief q (where $0 \geq p \leq 1$), in p, a person has is difficult, but Ramsey suggested that this could be inferred by discovering the odds at which a person would bet on *p*. Ramsey referred to degrees of belief so measured as betting quotients, and they measure a person's degree of belief in something at a specific time, without calling into question whether that belief is reasonable or not. The rationality of the belief can be determined by placing appropriate constraints to ensure coherence, continuity, and in line with Keynes' principle of indifference. Ramsey's subjective theory is similar to that of de Finetti (1937) and his influence is clear in von Neumann and Morgensten (1953). Keynes (2013) recognised that Ramsey was right to frame his logic of partial belief around his theory of betting quotients, saying in a review of Ramsey's paper that,

> "Ramsey argues, as against the view which I had put forward, that probability is concerned not with objective relations between propositions but (in some sense) with

[38] Ramsey gives the example of arsenic. If we believe that arsenic is poisonous, we will not, when the opportunity presents itself to drink it, in fact do so. Beliefs do not lead to action per se, but action when those beliefs are called upon.

> degrees of belief, and he succeeds in showing that the calculus of probabilities simply amounts to a set of rules for ensuring that the system of degrees of belief which we hold shall be a consistent system. Thus the calculus of probabilities belongs to formal logic. But the basis of our degrees of belief—or the a priori probabilities, as they used to be called—is part of our human outfit, perhaps given us merely by natural selection, analogous to our perceptions and our memories rather than to formal logic."

Keynes saw that Ramsey was right to see that probabilistic relations are not objective and that probability relates to our degrees of belief. There are not many "logics": so long as we are thinking, we are thinking in accordance with the rules of formal logic (Keynes 2013; Frege 1997; von Mises 1998), but, our degrees of belief are part of "human logic", our search for truth in a world of uncertainty.

Keynes Contra Induction

Keynes was unable, however, to fulfil his higher aim of arriving at a defence of inductive reasoning. Since Hume had shown the problem of induction, scientists and philosophers have not been able to develop a robust defence of inductive reasoning. It is clear that science progresses thanks to it, but the problem of induction opened the door to falsification, and since then, we have understood that there are no absolute truths, merely provisional truths (Lakatos 2015; Popper 1959). We have discussed the difficulties of testing an idea, such that in practice, verification is truer than falsification, and indeed, Lakatos (1965) argued that we advance by seeking to confirm rather than reject hypotheses, but the openness to falsification in itself separates science from non-science and ascience[39]. Scientific statements are, more precisely, fallible statements, statements that could be wrong. There are no crucial experiments, and no definitive proofs[40] or definitive falsifications, rather, scientists agree to accept as settled a certain hypothesis. This openness to being wrong has, since Hume, shown that settled notions may someday be proven wrong, or their

[39] One can say that scientific ideas are grounded in observable data, unscientific ones defy observable data, and ascientific ones are based on unobservable data. For instance, we can observe the elliptical movement of the earth, and if someone says the earth does not move, that is an unscientific idea, but if that person says that we live in a simulation, or that there is a spiritual world, that is an ascientific idea which we can believe or choose not to believe and which science cannot address (Hossenfelder 2022).

[40] Lakatos (1965) says, "Proofs do not prove, they improve".

boundaries severely changed, as happened when Newtonian physics had to adjust, first to relativity, and more starkly, to quantum physics. While Keynes failed to mount a definitive defence of induction, his notion that induction by analogy was more important than induction by frequencies, is powerful. For Keynes, the purpose of induction is to add weight to an argument. Here, he takes the same line as Lakatos in assuming that we seek to confirm rather than refute hypotheses. Arguments founded on analogy are based on the *likeness* of things, whereas arguments founded on frequencies are what he calls, “pure induction”. When we seek to limit the instances under which a hypothesis occurs, we use negative analogy. For instance, we could taste eggs, as Keynes proposes, in different parts of the year, and from different parts of the country, to see when eggs taste good and bad. Under pure induction, we do not know that we are increasing negative analogies. A positive analogy is one in which resemblances hold across board. Keynes finds that we cannot satisfy the condition that pure induction leads us to a point at which we achieve a finite certainty or uncertainty. Regardless of the number of instances we observe, we are still open to the possibility that we are wrong, and thus, we cannot make a truly cogent generalisation. Our inability to arrive at a finite probability presents us with a weak defence for pure induction.

The Art of Plausible Inference

The pre-Pascalian history of probability gives us a window into a grander view of probability and therefore, of risk. Probability is not a merely mathematical subject, even if, since the mathematization of probability, it has not been seen as a purely mathematical subject. The disputes between frequentist and Bayesian interpretations of logic can be overcome by the realisation that neither interpretation is valid everywhere and that the objects of study and sources of risk are not simply objects outside of us, but, our very thoughts and ideas. E.T. Jaynes (2003) argues that the dispute becomes irrelevant once we conceive of probability as extended logic.

Plausible reasoning, characterised by thinking of the form of equation (1), defines the kind of thinking that we undertake every day, and which humanity has undertaken since time immemorial. It can be thought of as an “inferior” model of thought to deductive reasoning, because it does not have the wealth of information that makes deduction possible (Jaynes 2003; Pólya 1968). The weaknesses of plausible reasoning are not a condemnation of it, but a feature of decision-making under uncertainty. One is apt to believe after a thousand years of data, that no black swans exist. Forecasting errors are commonplace, and, as Wucker

(2016) argues, even where we are able to predict the future, there are classes of high-impact threats that we ignore, or do not defend ourselves against with due seriousness, such as climate change. Although Taleb's discussion of black swans is popular, many of the examples he gives of black swans can be seen as one of Wucker's "grey rhinos": high-impact, predictable and yet ignored threats. For instance, many Austrian and post-Keynesian economists predicted the Great Recession of 2008 (Bezemer 2009), and despite many people referring to the Covid-19 pandemic as a black swan, Taleb has said it was predictable (Avishai 2020), and there is evidence that many experts predicted a pandemic[41]. These errors are a feature of plausible reasoning, where the certainty of deductive reasoning does not exist, and they simply highlight the epistemic nature of risk.

Despite the weaknesses of plausible reasoning, it remains a very powerful and very apt way of reasoning through uncertainty (Jaynes 2003; Pólya 1968). As early as 1662, Antoine Arnauld and Pierre Nicole attempted, through their *Logic: or the Art of Thinking*, to incorporate the mathematical theory of probability into a theory of logic, with their chapter on the application of mathematical ideas to the real world, written with the guidance of Pascal (Franklin 2015). Leibniz also attempted to develop a probabilistic logic, in his *New Essays Concerning the Human Understanding* in 1704 to aid in the determining the most plausible conclusion where it is not possible to arrive at a definite conclusion. Leibniz says that,

> "The Mathematicians of our day have begun to estimate chances in connection with gambling games. . . . I have more than once said that we should have a new kind of Logic which would treat of degrees of probability, since Aristotle in his Topics has done nothing less than that. . . . But he did not take the trouble to give us a necessary balance to weigh probabilities and to form solid judgements accordingly." (Leibniz 1916, 657)

Despite this, it has not been possible to apply mathematics to areas such as law, or come close to anything like an objective probability. This is clear from Leibniz' definition of a "fair game" as one where "there is the same proportion of hope to fear on either side" (de Melo and Cussens 2004, 33). In other words, there is the possibility of winning and losing on either side. The ratio between hope and fear need not be equal, but it must be constant, although Leibniz refers specifically to situations where there is an equiprobability of hope and fear, saying,

---

[41]The story of the virologist, Nathan Wolfe, who tried to sell a pandemic insurance package prior to the pandemic, is an apt example (Ratliff 2020).

> Axiom. If players do similar things in such a way that no distinction can be drawn between them, with the sole exception of the outcome, there is the same proportion of hope to fear. (de Melo and Cussens 2004, 33).[42]

Whereas Keynes' principle of indifference is epistemic and asks us to give even odds when we do not have any rational basis for asserting who will win, Leibniz takes this not as an epistemic but an aleatory condition, i.e. it arises not from ignorance but from the nature of things (de Melo and Cussens 2004).

However, in the most interesting aspects of life, where the value at risk is high, from law to science, there are no fair games. In law, "probability" is still not the probability of mathematicians (Franklin 2015). Moreover, as we have argued, the weighing of theories in science makes no use of a quantitative theory of probability. Franklin (2015) argues that neither Sir Isaac Newton nor Charles Darwin viewed probability as useful for weighing uncertain evidence and that even within mathematical literature, conjectures are not proved probabilistically. Artificial intelligence systems, for a practical example, are built on top of defeasible reasoning rather than a quantitative theory of probability (Strasser and Antonelli 2019; Donini et al. 1990; Franklin 2015).

Our discussion to this point has established that in navigating through risk, one is navigating through both measurable and immeasurable kinds of risk, through domains in which a frequentist approach seems rational and ones where such an approach is inappropriate, particularly in weighing the merits of different arguments, in the search for truth, and in in the face if highly complex situations, where the assignment of a number is purely communicative, as with the use of words of estimative probability.

Furthermore, we have shown that risk is not defined by its measurability, but by its tractability with either mathematical or logic tools, with the uncertainty of one's conclusion rising with complexity, a complexity which, at its height, reduces probabilities, if they are used, to merely communicative purposes. Thus, the risk of portfolio loss is as much a risk as the risk of choosing the wrong portfolio measure, or wrong portfolio theory, in as much as a legal system has the risk of unfair outcomes. Not all these can be enriched by the use of probabilities, and they may even be harmed by them.

We have also argued that de Finetti's insight that probability is subjective, that it "does not exist" is true, at least as uncertainty rises. We may argue that the probability of a heads is

[42] De Melo and Cussens (2004) discuss the similarities of this with Keynes' principle of indifference.

objective, but on the other end of the risk-uncertainty spectrum, which is arguably where the most interesting aspects of human action take place, that is not true, otherwise there would not be any uncertainty.

The subjectivity of probability is not a modern notion. We have seen that Bayes (1763) founded his theory of probability on expectation. This notion is present in Abraham Wald (1950), and de Finetti (2017; 1972), who both believed, despite their differences, that inference is impossible without value judgments. Our naive extension of Godel argues the same. In understanding that the fact-value dichotomy does not in fact exist, we can see that value judgments are implicated in our factual statements. There is no rain without the conceptual category of rain.

De Finetti's theory rests upon his idea of "coherence", where a gamble is assigned probabilities in such a way that the gambler is not a certain loser regardless of the outcome. Taleb (2018) posits a similar idea when he calls for "skin in the game" as the basis for fair decision making. Yet, our discussion shows that risk and probability extend beyond areas where one can make a bet. If a person is choosing between risk measures, the success of that portfolio may be in spite of that risk measure and not because of it, and so, the willingness to put money on the line, or create as-if scenarios, is not as fertile as first seems. Jaynes (2003) adds to that an aesthetic revulsion at basing theory on something as vulgar as gambling, and expectation. Furthermore, Jaynes shows that it is possible to justify the Bayesian approach on the work of Richard Threlkeld (Cox 1961; 1946).

Cox's theorem established probability as a general logic of plausible inference. Cox was able to derive the laws of probability, subject to the following desiderata: divisibility and comparability[43]; common sense (inc. adherence to Aristotelian logic)[44]; and consistency[45] postulates. Thus, objective Bayesianism, or logical probability is formed. It distinguishes itself from subjective Bayesianism in the following way: objective Bayesianism holds that if people have the same information, they will arrive at the same reasonable expectation, given that we all obey the same rules of logic. Subjective Bayesians believe that probability reflects personal beliefs, and that, therefore, all probabilities are equal.

[43] A proposition's plausibility is a real number that comes from knowledge we have about the proposition.

[44] There should be a sensible variance in the assessment of plausibilities.

[45] Derivations of the plausibility of a postulate are equal.

When we conduct plausible reasoning, we also assess degrees of plausibility, based on prior information collected from our experience (Jaynes 2003; Pólya 1968). Thus, even though we may err in our conclusions, they are rational based on what we know. The rationality of plausible reasoning does not equate to always arriving at the right conclusion. Jaynes (2003) argues that using plausible reasoning, when quantified, is able to arrive at the certainty of deductive reasoning. Indeed, mathematicians employ this weaker form of reasoning in their own work:

> “Good mathematicians see analogies between theorems; great mathematicians see analogies between Analogies.” Stefan Banach (quoted in Jaynes 2003, 6)

Cox’ theorem not only justifies a Bayesian approach to probability, it also allows us to conceive of probability as an extension of logic that guides decision making under uncertainty through the use of plausible reasoning (Cox 1946; 1961; Jaynes 2003; Van Horn 2003).

Part II: The Prospect of Ruin

The Pascalian Revolution

The Maximisation of Value

The revelation brought about by the Pascalian revolution came as a result of a fundamental break with prior thinkers: stochastic probability is made possible by the insight that the past determines, to a significant extent, the future (Bernstein 1996). When Gottfried von Leibniz wrote to Jacob Bernoulli in 1703 that, "[N]ature has established patterns originating in the return of events, but only for the most part" (Bernstein 1996, 4), his observation prompted Bernoulli to develop the Law of Large Numbers and methods of statistical sampling, that made it possible to get closer to an estimation of the patterns governing the world and therefore, to make better predictions about the future. Yet, Leibniz was sage enough to add, “for the most part”, understanding the presence of uncertainty due to the irreducible complexities and irregularities in nature: if the past was a perfect mirror of the future, nothing would ever change and human action would be impossible (Bernstein 1996; von Mises 1998). The notion that there is a deeper order behind randomness, was an insight that had escaped the Greeks, and the development of the mathematical theory of probability over a millenia since Aristotle had first laid the grounds for such a theory, was truly revolutionary.

The Problem of Points

The genealogy of expected value theory can be traced back to mathematician’s first attempts to develop a solution to the problem of points. The problem of points occurs when two players, with equal chances of winning future rounds, and who contribute equal amounts to winnings, and agree that the winner of a predetermined number of rounds will win the game, are forced to abandon their game before that predetermined number. Those players are then faced with a dilemma: how to divide the stakes fairly.

One of the earliest solutions to the problem of points was proposed by the monk Luca Pacioli in 1494 in his book, *Summa de arithmetica, geometrica, proportioni et proportionalità*, with other Renaissance mathematicians such as Niccolò Tartaglia and Giovan Francesco Peverone grasping for a solution, yet none was found (Edwards 1982; Devlin 2010; Bernstein 1996; Franklin 2015). Indeed, Tartaglia so despaired of finding a solution that he declared that any solution was “judicial rather than mathematical”, and would be so contestable as to lead to litigation. Part of the difficulty in finding a resolution to the problem is that dice throwing is not inherently theoretically interesting, even if people have, as we have seen, been throwing dice since time immemorial. Not only is the subject terribly tedious, it is not directly useful: one does not become a better gambler simply because one understands the probability distribution of a throw of dice.

It was not until the late summer of 1654 when the chevalier de Méré presented the problem to Blaise Pascal, that a solution was found (Edwards 1982; Devlin 2010). Out of the problem of points, the notion of expected value was born. Pascal presented his solution to Pierre de Fermat, stating that he had discovered a way to calculate a fair division of stakes, later providing his complete solution in his book, *Traité du triangle arithmétique avec quelques autres petits traitez sur la mesme matière*. The correspondence between the two mathematicians, in which they developed three solutions to the problem, reveals that they understood as a fundamental principle that a fair division of stakes had to reflect the chance of getting it.

Having seen Pascal's (1665) solution, Christiaan Huygens' book, *De Ratiociniis in Ludo* examined the problem of points and presented Pascal's expectations-based solution to the world[46] in the world's first systematic treatise on the mathematical theory of probability. More than a millennia since Aristotle had laid the foundations of a mathematical theory of probability, one had finally emerged. Huygens established the notion of expected value[47], saying,

> That any one Chance or Expectation to win any thing is worth just such a Sum, as wou'd procure in the same Chance and Expectation at a fair Lay. As for Example, if

[46] Huygens may have come upon the same solution as Pascal, or, more likely, it was communicated to him prior to his book's publication (Edwards 1982).

[47] The expected value E[X] of a finite set of random variables, X, is the weighted average of all possible outcomes, emerged from attempts to resolve the problem of points.

Although Huygens' (1714) introduced the term "expectation" to the lexicon of probability theory, it was not until Pierre-Simon Laplace's 1814 tract, *Théorie analytique des probabilités*, that expected value was formally defined.

> The probability of events may be useful in determining the hope or fear of people affected by their occurrence. The word expectation has various interpretations: in general it expresses the advantage to anyone who expects any benefit whatsoever, under assumptions that are only probable. This advantage, in the theory of chances, is the product of the sum expected and the probability of getting it; it is the partial sum that ought to be paid out when one does not want to run the risks of the event, supposing that the total sum is to be shared out in proportion to the probabilities. This allocation is the only fair one when all irrelevant circumstances are set aside, because an equal degree of probability gives an equal claim to the sum expected. We shall call this advantage *mathematical expectation*. (Laplace 1995, 11)

> any one shou'd put 3 Shillings in one Hand, without telling me know which, and 7 in the other, and give me Choice of either of them; I say, it is the same thing as if he shou'd give me 5 Shillings; because with 5 Shillings I can, at a fair Lay, procure the same even Chance or Expectation to win 3 or 7 Shillings.

In Huygens game, as we shall call it, we are presented with a risky proposition, X, with a set of possible payouts, $x_i \in \{3, 7\}$, with associated probability distribution, $P(x) = 0.5 \ \forall \ x \in X$. The expected value of the game is defined by,

$$E[X] = \sum_{i}^{n} x_i p_i \tag{3}$$

(3) is of course the finite ensemble average of the gamble. With an E[X] of 5, Huygens makes the claim that this game "is the same thing as if" he was given 5 shillings. The economic agent, assuming initial wealth w_0, will, of course, look to maximise expected value, which is to say that they will seek the largest profit, Π, possible, where

$$\Pi = \sum_{i}^{n} x_i p_i - x_0 \tag{4}$$

Huygens' belief that the game amounts to being given the expected value is of course obviously both true and untrue. For instance, there is a material difference if one picks the hand with 3 shillings over the hand with just 7 shillings. The truth of Huygens' statement rests on the dynamics of the game as we repeat it. If we played Huygens' game 100 times, the mean payout might be 4.8. Per the central limit theorem, as $i \to \infty$, the mean payout approaches the expected value. The more we play, the closer we get to the *limit* of the game, i.e. the expected value. If we play this game 1 million times, the mean payout might be 4.999444.

$$E[X] = \lim_{i \to \infty} \sum_{i}^{n} x_i p_i \tag{5}$$

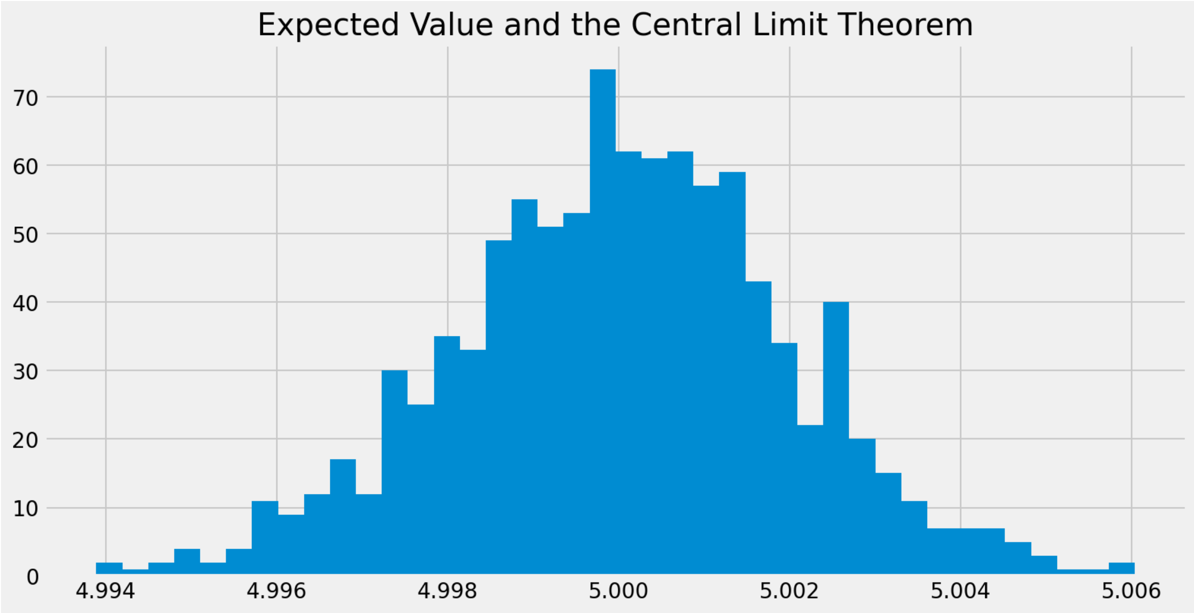

Figure 2 : Expected Value and the Central Limit Theorem

However, Huygens' gamble is unfair: one wins whether it's heads or tails, the left hand or the right hand. To truly appreciate the meaning of expected value and its impact on wealth, we propose the following experiment: assume initial wealth, w_0 = $5,000, which you will stake on a gamble, X, defined by possible outcomes $x_i \in$ {0.6*5000, 1.40*5,000} with associated probability distribution, $P(x) = 0.5\ \forall\ x \in X$, and played out over i = 100 iterations. The expected value of the game is

$$E[X] = 0.5(0.6*5000) + 0.5(1.4*5000) = 5000 \quad (6)$$

If we played this game 100 times, we could enjoy a mean payout of $5,120. If we played this game a million times, we could enjoy a mean payout of $4997.9. Once again, we observe that the more times we play the game, the closer the mean payout gets to its limit, E[X].

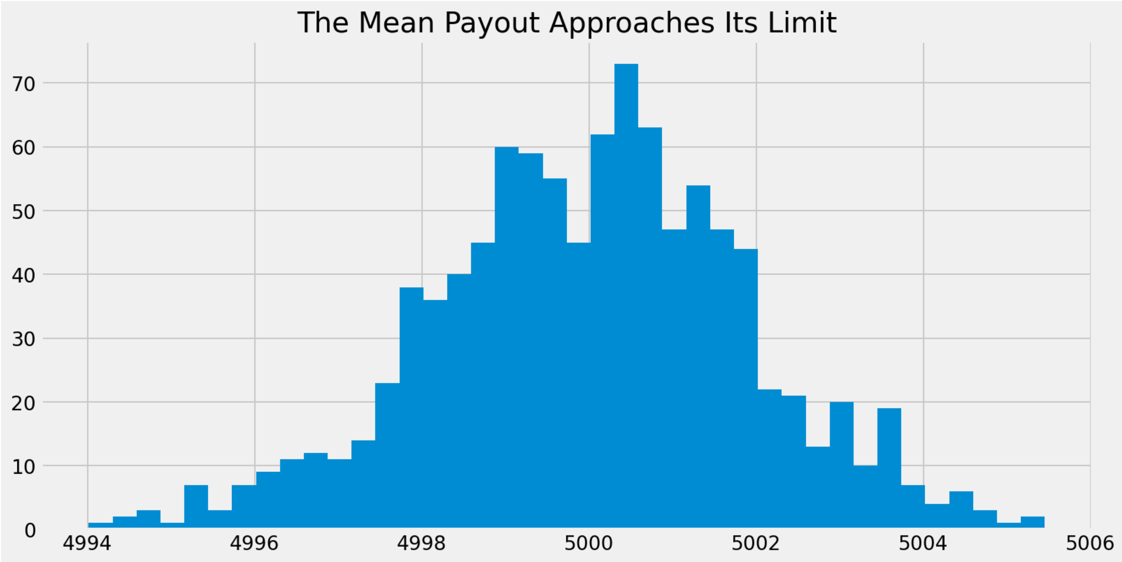

Figure 3 : The Mean Payout Approaches Its Limit

The Weaknesses of Expected Value Maximisation

Today, the maximisation of expected value is the dominant decision theory in the *practice* of finance. Managers are not only urged to maximise shareholder value (Damodaran 2014; Koller, Goedhart, and Wessels 2020), but it is assumed that such behaviour is rational, or more aptly, rationally obligatory (Pollock 1983). Nonetheless, human action seems to defy expected value maximisation. Consider the following scenarios: a manager is faced with two projects: project A and project B. Project A will generate a certain €3 million and therefore has an expected value of €3 million. Project B has two equi-probable outcomes: with a 50% probability, the project will generate €8 million and with a 50% probability, the project will generate nothing. Project B has an expected value of €4 million. Such “A/B choices” (Pollock 2010) are replete in real-world scenarios, between events of which one has a large degree of confidence, and ones in which one has rather less. Expected value maximisation theory tells us that picking B is rationally obligatory. A more cogent example of expected value theory highlights its pitfalls: a homeowner living just above the poverty threshold is presented with an enticing gamble by a “malevolent decision theorist”: she can risk her $80,000 home for a coin toss, which at heads, will earn her $85,000 tax free, and at tails, will cause her to lose her home. The expected value of this gamble is, of course, $82,500. The rationally obligatory decision is, according to expected value theory, to take the gamble. Yet, as John L. Pollock (2010) notes, few believers in expected value theory would take this gamble, instead choosing to segregate what they ought to do from what they would actually do, given their particular risk profile. One is left either with a notion of human thinking as irrational, or, we must question the prevailing doctrine. The maximisation of expected value is, to borrow from Gottlob Frege (1997), not true but simply regarded as true.

Daniel Kahneman and Amos Tversky (1979) have shown that people often prefer a certain gain to a prospective one[48], and that losses weigh more heavily than gains in estimations of risky propositions[49]. For instance, McKinsey & Company surveyed 1,500 executives from 90 countries and assessed their reactions to different investment scenarios (Koller, Lovallo, and Williams 2012). They found that managers displayed “extreme levels of risk aversion” even when faced with projects whose expected value was “strongly positive”. Managers were unwilling to accept projects with a positive expected net present value with a risk of loss of 75%. Managers were only willing to accept a risk of loss of between 1% and 20%, regardless of the size of the investment. McKinsey & Company, of course, urged managers to “overcome their bias against risk” and be risk-neutral, or hijack their narrow framing by pooling risks.

Huygen’s Criterion Leads to Gambler’s Ruin

The rupture between theory and practice can be understood through the framework of ergodicity, which, as we will see later, is at the heart of the Saint Petersburg Paradox (Peters 2011; 2019; 2020; Peters and Gell-Mann 2016). Economic theory assumes that the time average and ensemble average are equivalent, which is to say, that what happens across time is equivalent to what happens when time is not taken into account. This ergodic hypothesis is true in equilibrium systems, but much of economic theory deals with far-from-equilibrium systems (Peters 2011; 2019; 2020; Peters and Gell-Mann 2016). In other words, there is an appreciable difference in how objects behave when there is no time involved, and where there is time.

Thus, Huygens drew the wrong conclusion. This was inevitable because the mathematical theory of risk emerged from games of chance, which are a poor analogy to the field of action through which people make decisions, where each economic agent has different levels of prior wealth, takes a different position on a trade, and invest different shares of their prior wealth, into a risky proposition. In addition, Huygens’ game essentially plays out in parallel universes, from which he draws a mean payout, but economic agents do not draw a mean payout from parallel universes, rather, they draw payouts across time. When we introduce time, Huygens’ gamble becomes disastrous.

48 The certainty effect.
49 Loss aversion.

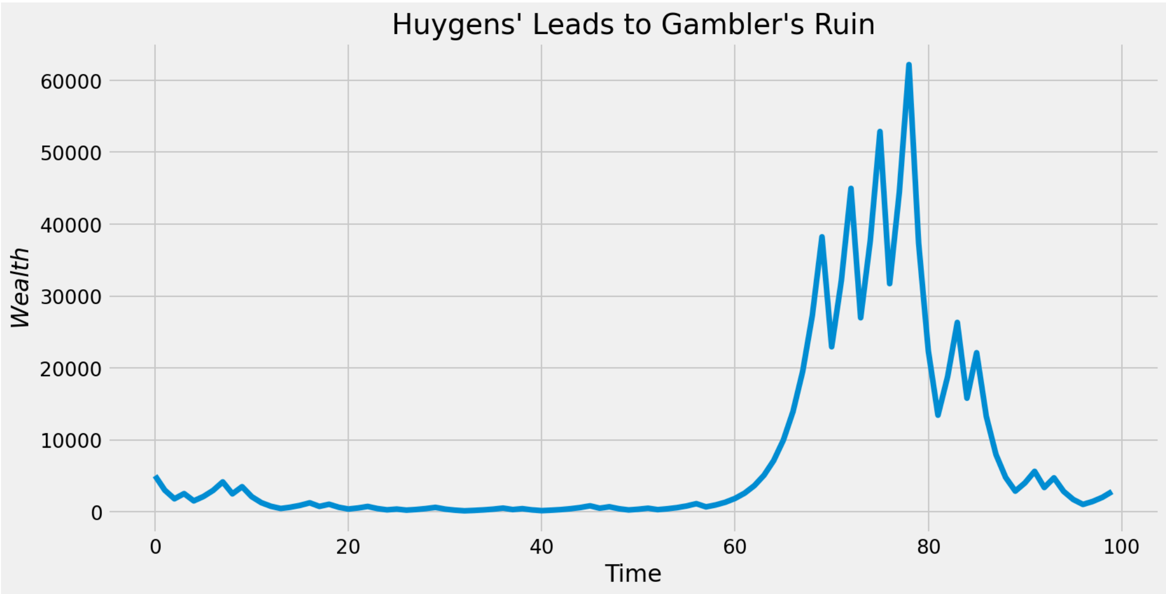

Figure 4 : Huygens' Leads to Gambler's Ruin

Gambler's ruin is the ultimate result of this gamble. Whereas Huygens' gamble assumes that wealth is ergodic[50], it is actually non-ergodic (Peters 2011; 2019; 2020; Peters and Gell-Mann 2016). Perhaps one was unlucky? We could, like a cat, play this game across different lives. Nevertheless, it becomes clear that though one may enjoy periods of luck, when fortune seems to smile on one, but in the long run, gambler's ruin is certain.

[50] Ergodicity refers to a state where the ensemble average (played out across parallel universes) is equal to what you would get across time, the time average (Peters 2019; Shannon 1971).

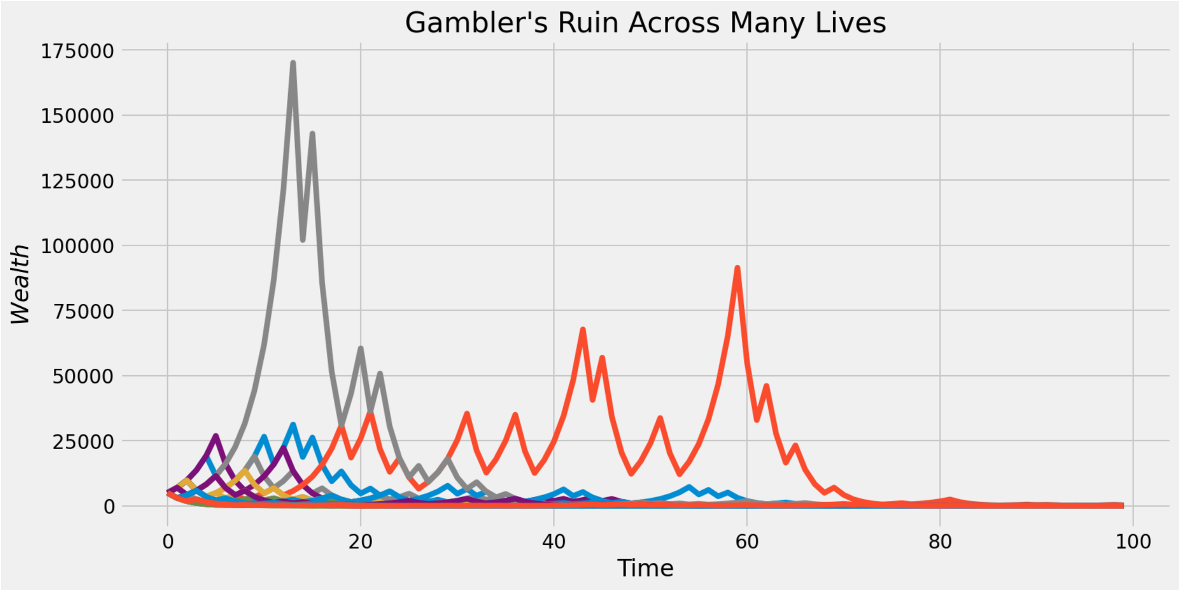

Figure 5: Gambler's Ruin Across Many Lives

As $t \to \infty$, wealth's random walk smoothes out and terminal wealth, $w_t \to 0$.

$$\lim_{t \to \infty} w(t) = 0.6^{t/2} 1.4^{t/2} = 0 \qquad (7)$$

The time average of the gamble is 0 despite it possessing a positive expected value. In a game in which one should come out ahead, according to Huygens' decision criterion, wealth is instead wiped out. Huygens presents us with a static notion of risk that ignores the fact that gambles occur over time, and that there is a fundamental difference between averaging the result over many trials, and averaging over time (Peters 2020). In real life, rather than *i* iterations run in parallel, we have only one iteration run across δt intervals of time *t*. The impact of this is clear: wealth grows slower, or decays faster over time than across an ensemble, indeed, ensemble wealth can grow while individual wealth decays[51]. The expected value is not equal to the time average.

The Pascal-Huygens' decision criterion leaves one with two options when faced with a risky proposition: the trivial gamble of non-participation, or, selecting one of the risky propositions on offer and attempting to maximise wealth (Pollock 1983; Peters 2020). Maximising expected value implies that one is rationally obligated to accept a risky proposition, X, if E[**X**]

[51] In Misesian language, knowing how the class of events behaves is irrelevant to the fortunes of a single event. A class of bets may, on average, earn a profit, but in the real world, one does not draw the class average.

> 0 and E[**X**] > $E[\mathbf{X}]_A$, where $E[\mathbf{X}]_A$ is the expected value of an alternative risky proposition. This is extendable over more than one alternative risky proposition, to include the set, S, or alternative risky propositions, so that one is rationally obliged to select a risky proposition *iff* E[**X**] > $E[\mathbf{X}]_S$ (Pollock 1983). However, this simple principle is problematic. Defining what counts as an "alternative" risky proposition is harder than first appears. For instance, we may consider as rationally obligatory, any risky proposition whose expected value is greater than that of any risky proposition incompatible with it, and as rationally permissible any risky proposition as least as great as any incompatible risky proposition[52]. Suppose that you are given three risky propositions, A, B and C, where A and B are incompatible with each other and thus, participating in one precludes participation in the other, and C can be taken up independently of both, and E[A] = E[B], and that by participating in B and C, E[BC] = E[B] + E[C]. Then, it is rationally permissible to take up A or B, and that BC is both rationally obligatory and incompatible with A, so that one is rationally obliged to refrain from A. Pollock (1983) argues that even in trying to define expectation in terms of the expectation of refraining from an act, we encounter difficulties. He suggests that the most realistic ways to define it in the following way:

An act A is rationally obligatory iff it is prescribed by a maximal strategy[53] that is rationally preferable to one which does not prescribe it, where a maximal strategy is deemed rationally preferable if it has a higher expected value than the second. (8)

Which can be reaffirmed and strengthened as,

Given time t_o, a maximal strategy S is rationally preferable to another S' iff $E(S_t) > E(S'_t)$ for all t_n, where $n > 0$. (9)

It is often argued that to escape the supposed narrow framing of economic agents, they must think in terms of a portfolio, rather than in terms of individual bets or projects (Koller, Goedhart, and Wessels 2020; Koller, Lovallo, and Williams 2012). However, this approach still carries many of the same risks that drive actors toward risk aversion: the accumulation of potentially existentially threatening risks in a world where everything that can go wrong, may go wrong.

[52] Definitions adopted from (Pollock 1983), who refers more broadly to "acts".

[53] The expectation value of a strategy, according to Pollock (2010), is the expected value of undertaking to conform to a strategy rather than of actually conforming to it (which would exclude failed attempts at conforming to it).

Pollock (2010) strikes upon a very Bernoullian ideal in his examination of expected value theory. He starts from the “situation value” of the actor, in other words, the utility value of a person’s accumulated goods, or prior wealth before a risky proposition. To use language more in keeping with this treatise, we have to consider the risk of a proposition from a subjective point of view, reflected in that person’s prior wealth. Assuming a linear wealth function, rather than one reflecting decreasing marginal returns, Pollock shows that a person’s situation value rapidly becomes catastrophic below a poverty threshold. Once a person has fallen into “the pit”, it is difficult for that person to get out. In his scenario, a person has a 1% probability of dying in any period of time, termed a “cycle”. The poverty threshold is 100. Wealth, G, decays by 7 units a cycle, with a floor of 1 unit. A person has a 9% chance of being offered an A/B choice, which is the only way that they can build wealth. A gives a certain 300 units, B comes with two outcomes: a 50% chance of gaining 800 units, and a 5% chance of getting nothing. The expected value of B is 400.

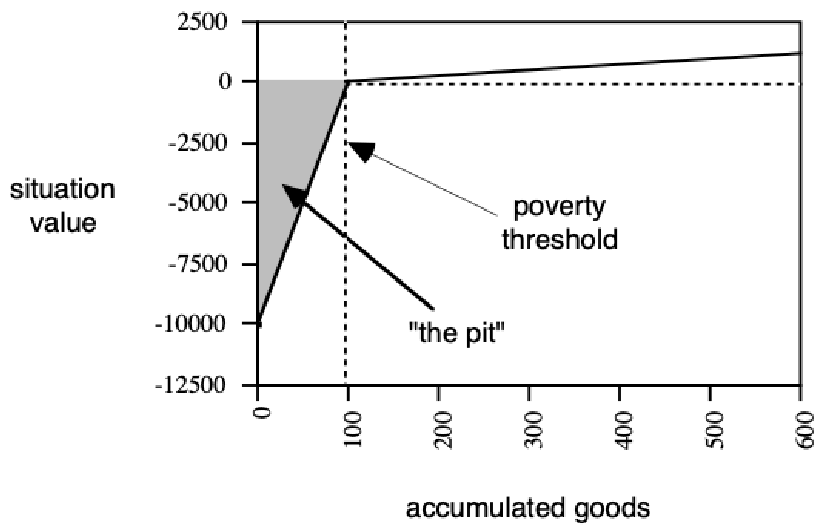

Figure 6 : Situation-value as a function of accumulated goods. Source: Pollock (2010)

Rather than selecting B, he suggests that a person is faced with three choices: choosing A, which is sub-optimal according to expected value theory, or B, which maximises expected value, or a third option, C, where B is chosen iff wealth is 200 units, in other words, C recommends risk aversion when a person is poor, and maximising expected value when a person is rich. He finds that the success of A, B or C is dependent upon a person’s prior wealth. Maximising expected utility is a bad strategy for all but the wealthy. Where a person has less than 200 units of G, banking on a sure thing dominates maximising expected value.

Over 200 units and a policy of maximising expected value dominates a policy of seeking a sure-thing. C dominates both strategies. Risk aversion is, thus, rational for an impoverished person, the degree of risk aversion rising with the size of a wager. As a person accumulates wealth, they can take on more risk.

Figure 7 : Life-accumulations. Source: Pollock (2010)

The impoverished, or resource constrained actor cannot tolerate the losses that a more affluent can, because of the risk of falling into “the pit”, out of which it is very difficult to emerge.

The Saint Petersburg Paradox

After Huygens established expected value as the fair approach to valuing risky propositions, it was assumed that, faced with different risky propositions, agents would select the one that results in the greatest increase in their wealth. Nevertheless, over time, the settled notion of expected value became increasingly unsettled. Nicolas Bernoulli questioned the merits of expectation, showing its fallibility through what became the archetypical paradox of the era: Saint Petersburg paradox, which he invented in a letter to Pierre Rémond de Montmort (de Montmort 1713). Originally framed with a single die, Nicholas Bernoulli’s cousin, Daniel Bernoulli (1954), turned it into a coin-flipping game:

> Peter tosses a coin and continues to do so until it should land "heads" when it comes to the ground. He agrees to give Paul one ducat if he gets "heads" on the very first throw, two ducats if he gets it on the second, four if on the third, eight if on the fourth, and so on, so that with each additional throw the number of ducats he must pay is doubled. Suppose we seek to determine the value of Paul's expectation.

Nicholas Bernoulli suggested it would be “curious” to calculate the E[X] of the risky proposition.

$$E[X] = (\tfrac{1}{2} \times 1) + (\tfrac{1}{4} \times 2) + (\tfrac{1}{8} \times 4) + (1/16 \times 8) + \ldots = \tfrac{1}{2} + \tfrac{1}{2} + \tfrac{1}{2} + \tfrac{1}{2} + \ldots = \infty \qquad (10)$$

Daniel Bernoulli observed that,

> Although the standard calculation shows that the value of Paul's expectation is infinitely great, it has ... to be admitted that any fairly reasonable man would sell his chance, with great pleasures, for twenty ducats.

Whereas, in his *Pensées*, Pascal had argued that, in face of infinite grace, a person must be prepared to pay anything, the Saint Petersburg paradox shows us that, given a risky proposition with a potentially infinite payoff, the typical person could hardly be persuaded to pay a fee of no more than “twenty ducats”. Infinite wealth, it seems, is worth less than infinite grace. The paradox exists in part because real-world decision making was not taken as irrational, rather, theory had to answer why it diverged so sharply from human action.

There are practical objections to the Saint Petersburg paradox: Alexis Fontaine des Bertins, for example, noted that any casino or banker in the real world would not have the infinite resources needed to come good on their end of the game and a gambler would know this (Dutka 1988). Therefore, it would not make sense for anyone to pay a large fee for a ticket that was likely unredeemable. In addition, no casino or banker would offer a lottery with an infinite expected loss.

However, these practical objections miss the point. The Saint Petersburg paradox retains its power even with a more limited payoff. The infinite payoff only dramatises a deeper lesson. If, rather than having a series that grows without bounds, we cap the wager at $1 billion, we would still observe the same features of the wager (Peters 2011; Spitznagel 2021; Poundstone 2005).

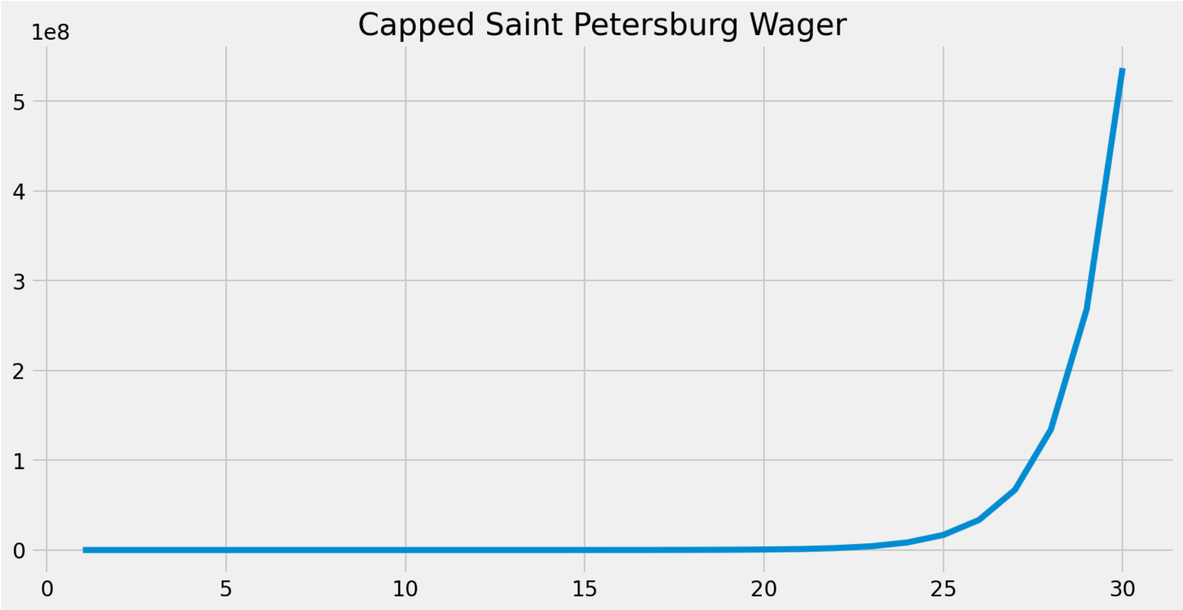

Figure 8[54]: Capped Saint Petersburg Wager

At the 31st flip of the coin, the payoff would be $1,073,741,824, so, the number of tosses should be capped at 30, with the winnings awarded to the player who gets to the 30th tosses. The E[X] of the wager would drop from ∞ to $15, an E[X] with a probability of 0.00000009313225746%. A sensible person would not, it seems, gamble heavily on an outcome with vanishingly small odds, no matter how positively skewed the payoff distribution. Indeed, it must be said that a player who paid $15 to play this capped Saint Petersburg wager would only receive a payoff greater than the value of the fee on the fifth toss. Expressed this way, the paradox seems facile: of course people would be reluctant to pay a large fee when earning the ultimate payoff was extremely unlikely and where they had to survive several rounds of losses just to breakeven! Bernoulli points out that,

> it appears that in many games, even those that are absolutely fair, both of the players may expect to suffer a loss; indeed this is Nature's admonition to avoid the dice altogether....

In reality, if we played this game 30 times, we might have a mean payout of just over 491 ducats. Played a million times, the mean payout might shoot up to nearly 499,862 ducats. The longer we play, the closer the mean payout would get to expected value: ∞.

[54] See data in the appendix.

However, Bernoulli goes beyond this. He argued that expectation was an “irrational” approach to financial decisions. Of that, everyone is in agreement. Orthodoxy tells us that his paper resolved the Saint Petersburg paradox by pioneering expected utility theory[55]. A close reading of the text suggests otherwise.

Daniel Bernoulli’s Theory of the Maximisation of Geometric Returns

The Subjective Basis of Risk

Bernoulli points out that expected value theory rests upon the presumption of the objectivity of risk, that is to say, faced with the same risky proposition, a coin toss or throw of the dice, two agents face the same risks and consequently, the same expected value. It can therefore be implied that they will make similar decisions. In other words, a changing cast of players facing the same risky proposition, should all make the same decisions. Risk is, therefore, a question of rules, rather than judgement, and all that is left is

> the enumeration of all alternatives, their breakdown into equi-probable cases and, finally, their insertion into corresponding classifications.

A person's subjective characteristics are, so went standard discourse, irrelevant to the calculation of risk. However, Bernoulli showed the inherent subjectiveness of the drivers of risk, by highlighting the importance of prior wealth in the estimation of risk. Therefore, two people faced with the same risky proposition, have, rather than the same risks, different risks and therefore, different estimations of the value of any gain from that proposition. Bernoulli (1954, 24) says the value of the risky proposition must be re-established and that,

> To do this the determination of the value of an item must not be based on its price, but rather on the utility it yields.

His use of the word "yield" is interesting because it suggests a *rate of return*. The importance of this will become clearer, later. Bernoulli goes on to say that,

---

[55] Expected utility can be traced to Gabriel Cramer's solution to the Saint Petersburg Paradox and Bernoulli (1954), both of which sought to rebuild value on a new basis, after the expected value theory of Pierre, Fermat and Huygens had seen to run aground.

> The price of the item is dependent only on the thing itself and is equal for everyone; the utility, however, is dependent on the particular circumstances of the person making the estimate.

Having stated his belief in the subjective basis of risk, Bernoulli restates his theory of value as follows,

> If the utility of each possible profit expectation is multiplied by the number of ways in which it can occur, and we then divide the sum of these products by the total number of possible cases, a mean utility [moral expectation] will be obtained, and the profit which corresponds to this utility will equal the value of the risk in question.

Grounded on the "particular circumstances" of the individual, utility is therefore, changeable with the change in circumstances, and rational agents should maximise their utility.

Utility Derives from Growth

Louise Sommer, who, with the assistance of Karl Menger, translated Bernoulli's paper into English, centuries after it first appeared in Latin, renders the Latin term, "*emolumentum medium*", as "mean utility" (or even "moral expectation", borrowed from George Cramer), but the term would have been more properly translated as mean or average "advantage", "benefit" or "profit". The difference between Bernoulli's vision of utility and the modern interpretation of utility is different enough that I will use the term, "Bernoulli utility" or, “EM”, to distinguish the two. Indeed, Ole Peters (2020) argues that Bernoulli’s expected utility is not even mathematically equivalent to von Neumann and Morgenstern’s theory of utility.

Long before the Marginal Revolution of Carl Menger, Léon Walras and William Stanley Jevons, which established the subjective basis of value in modern economics, Bernoulli (1954, 25) opined that,

> Now it is highly probable that any increase in wealth, no matter how insignificant, will always result in an increase in utility which is inversely proportionate to the quantity of goods already possessed.

In other words, the utility derived from any increase in wealth is dependent on the level of one's prior wealth. The law of diminishing marginal utility. However, Bernoulli's vision of utility diverges from that of the marginalists: his utility is a function of prior wealth and not a consequence of a more psychologically grounded sense of "usefulness" of wealth. In reading Bernoulli closely, one can even make the argument that utility was a pedagogical device meant to lead the reader toward a greater insight, an insight not anchored in the psychologism of the marginalians or von Neumann and Morgenstern (1953). For Bernoulli, utility comes from a growth in prior wealth, it is about returns. Bernoulli cares about the percentage change in wealth. More importantly, in discussing inverse proportionality[56] Bernoulli is already hinting at his result.

In his Nobel Prize Lecture, Daniel Kahneman (2003) came close to realising Bernoulli's insight. He recognised that perception is "reference-dependent", and noted that this is incompatible with expected utility theory, tracing the problem to Bernoulli. Kahneman spoke for most economists in saying that the expected utility theory is the dominant model in academic economics, and that maximisation of expected utility was not only prescriptive, it also reflects how reasonable people behave. Kahneman believed Bernoulli's conflation of prescriptive behaviour and actual behaviour was an error. Kahneman accused Bernoulli of developing a reference-independent theory. This error motivated Kahneman & Tversky (1979) to develop prospect theory. However, looking at Bernoulli's model suggests that his model may have been flawed, but his instincts were very different from expected utility theory, and very profound.

Bernoulli's Model

Bernoulli demonstrates his theory with the following curve:

[56] Utility is frequently denoted by the function $u(x) = \log(x)$.

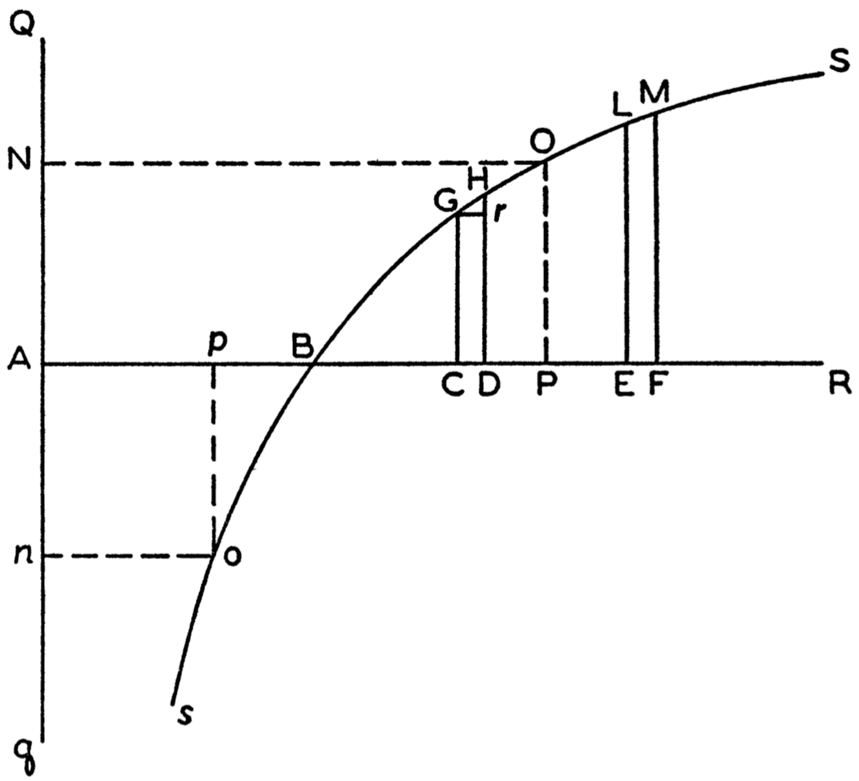

Figure 9: The Logarithm of Wealth. Source: Bernoulli (1954)

The geometric notation[57] can be translated as follows:

Metric	**Bernoulli's Notation**	**My notation[58]**
Prior wealth	AB	w_o
Initial gain	BC	π_i
Second gain	BD	π_2
Third gain	BE	π_3
Fourth gain	BF	π_4
Initial terminal wealth value	AC	$w_0+ \pi_1$
Second terminal wealth	AD	$w_0+ \pi_2$

[57] See section 6 of Bernoulli (1954).

[58] Adapted from Peters (2020).

value		
Third terminal wealth value	AE	$w_0 + \pi_3$
Fourth terminal wealth value	AF	$w_0 + \pi_4$
Utility of prior wealth	0	$\Delta u(w_0)$
Utility of π_i	CG	Δu^+_1
Utility of π_2	DH	Δu^+_2
Utility of π_3	EL	Δu^+_3
Utility of π_4	FM	Δu^+_4
Infinitesimally small gain in utility for a person with wealth $w_o + \pi_1$	rH, or dy	Δy
Infinitesimally small gain in wealth for a person with wealth $w_o + \pi_1$	CD, or dx	Δx
Probability of π_1	$m/(m + n + p + q + ...)$	p_1
Probability of π_2	$n/(m + n + p + q + ...)$	p_2
Probability of π_3	$p/(m + n + p + q + ...)$	p_3
Probability of π_4	$q/(m + n + p + q + ...)$	p_4
Emolumentum medium/Expected Bernoulli utility	PO=AN	EM
Bernoulli utility from loss	po=An	Δu^-
Maximum riskable wealth	Bp	R_m
Wealth associated with Bernoulli utility	AP	W_B

Expected gain in wealth	NO-AB	$W_B - w_0$

Table 1: Converting Bernoulli's Notation

The following equations can then be derived:

$$EM = \sum p_i \Delta u_i \tag{11}$$

Bernoulli is not interested in simply determining EM, he wants to know what level of wealth is associated with that utility, and so he proceeds past this introduction to utility theory, to what it seems he is really after, the corresponding expected gain associated with the risky proposition. Utility for Bernoulli is only a gateway to his notion of the "gain which may be properly expected", which is the difference between the wealth associated with the "moral expectation", and prior wealth.

$$W_B - w_0 = \Pi \tag{12}$$

The value of the risky proposition can be restated as,

$$\Pi = \Sigma p_i \pi_i \tag{13}$$

Estimating the expected gain or Bernoulli expected wealth is the true goal of Bernoulli's work.

In Corollary I, Bernoulli again restates conventional wisdom: expected gains are absolute, rather than relative, and the utility derived is proportional to the gain. There is the air of the mathematician and rhetorician in Bernoulli's work, and, through guided discovery, he hopes to lead his readers to his vision. Consequently, he does not state in advance of his revelation, where he is going, or what he is doing. In fact, he does not even state the obvious fact that the curve he is using is a logarithm. Instead, he guides the reader to this point. He points out that if this hypothesis, that expected gains are absolute and their associated utility directly proportional to them, the utility curve would in fact be a straight line, with (11) and (12) being true "in conformity with the usually accepted rule". This of course is distinct from his point that utility is inversely proportional to gains.

In Corollary II, he investigates "the nature of the curve" describing *emolumentum*, and reveals that it is, instead of being a straight line, it is the logarithmic curve of wealth.

Let $w_0 + \pi_1 = x$, $\Delta u^{+}_{1} = y$, $w_o = \alpha$, *b* be a constant, then,

$$\Delta y = b\Delta x \div x \tag{14}$$

The change in *emolumentum* is directly proportional to the change in wealth scaled by prior wealth. Which leads us to,

$$y = blog\,(x \div \alpha) \tag{15}$$

Thus, the curve is the logarithmic curve of wealth.

Bernoulli returns to (11) and finds that EM,

$$EM = b\log\,(W_B/w_o) \tag{16}$$

EM therefore, is directly proportional to the logarithm of wealth. Expected utility theory argues that we maximise for (17). However, Bernoulli's theory goes beyond that.

Possible terminal wealth values can be expressed similarly to (16),

$$b\log\,(W_B/w_o) = b^*\Sigma p_i \log\,[(w_0 + \pi_i)/w_o] \tag{17}$$

We can see that EM is the probability weighted average of the logarithms of all possible terminal wealth outcomes, scaled by prior wealth.

$$b\log\,(W_B) = b^*\Sigma p_i \log\,[(w_0 + \pi_i)] \tag{18}$$

This leads us to the following result,

$W_B = e^{EM}$ (19)[59]

We can then determine that

$$\Pi = W_B - w_o \tag{20}$$

[59] This result follows from but is not present in Bernoulli (1954).

W_B can be expressed as a geometric mean of possible terminal wealth values,

$$W_B = (w_0 + \pi_1)^{p1}(w_0 + \pi_2)^{p2}(w_0 + \pi_3)^{p3}...(w_0 + \pi_i)^{pi})^{1/100} \quad (21)$$

Bernoulli's insight is that wealth is multiplicative over a time-series. Whereas expected utility theory, and indeed, prospect theory, has been built on theories of human behaviour and motivation, Bernoulli has built his theory purely on an understanding of the need to multiply wealth over time. His theory not only rejects expected value maximisation, it also rejects the use -still widespread today-, of the arithmetic mean.

Bernoulli's goal was not to establish a theory of utility, his goal went beyond that, to determine the value of a risky proposition and in getting there, he utilised the geometric average. In so doing, he established the true decision criterion used by rational actors, saying,

> "Any gain must be added to the fortune previously possessed, then this sum must be raised to the power given by the number of possible ways in which the gain may be obtained; these terms should then be multiplied together. Then of this product a root must be extracted the degree of which is given by the number of all possible cases, and finally the value of the initial possessions must be subtracted therefrom; what then remains indicates the value of the risky proposition in question."

Nature's Admonishment to Avoid the Dice

In §13, Bernoulli warns investors that, given the "concavity of the curve", they face perilous hazards, and "that in many games, even those that are absolutely fair[60], both of the players may expect to suffer a loss; indeed this is Nature's admonishment to avoid the dice altogether". One faces gambler's ruin, and in the face of a risky proposition, non-participation is the safest course.

[60] This means the absence of advantage on either side, resulting in both sides having an expected value of the bet equal to zero.

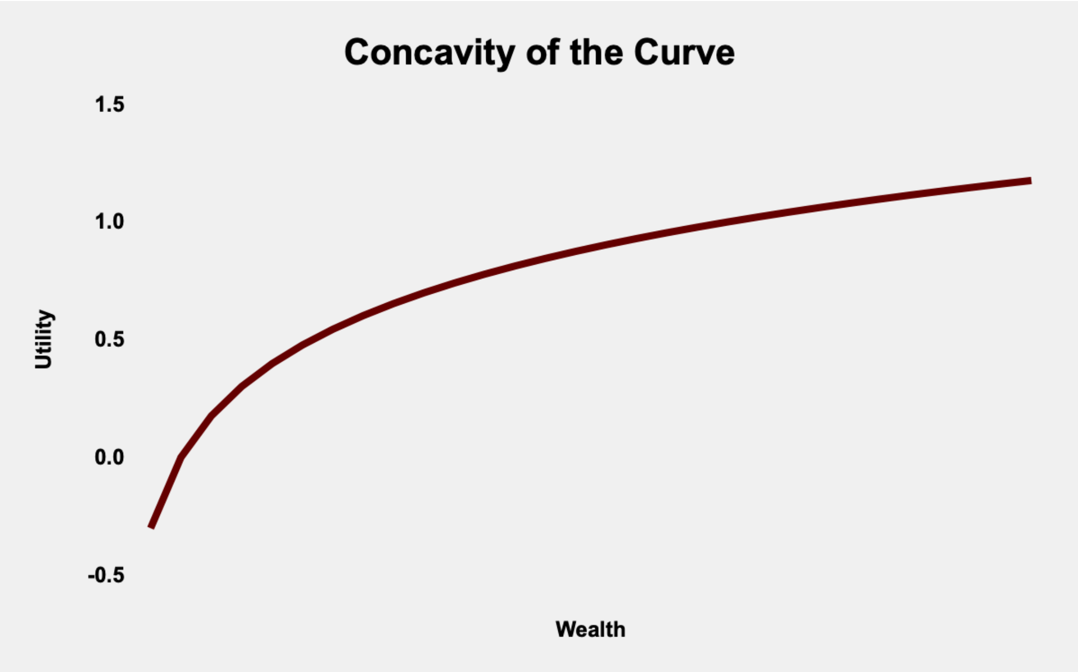

Figure 10: Concavity of the Curve

If $R_m = \pi$, then the disutility of any losses is disproportionately greater than expected gains of the same absolute magnitude. This is clear from a non-rigorous example: suppose two investors have the opportunity to invest in the initial public offering (IPO) of a Software-as-a-Service (SaaS) company, “Cloud-Cuckoo Land Inc.” whose IPO price will be 100 livres a share[61]. There are two possible outcomes for the next 12 months: a 50% rise in the share price, which will take the price to 150 livres, or a 50% decline in the share price, which will take it to 50 livres. Each outcome, IPO_{Gain} and IPO_{Loss}, has a 50% probability, resulting in the set of outcomes,

$$IPO = \{(50, 50\%), (150, 50\%)\} \quad (22)$$

Given (22),

$$W_B = (50^{50} * 150^{50})^{1/100} \approx 86{,}603 \text{ livres.} \quad (23)$$

Both investors will experience a loss of over 13 livres, even under fair conditions. Explaining the loss is simple: having suffered IPO_{Loss}, the intrepid investor would need a 100% subsequent increase in the share price for the stock to get back to $100. One is punished for losses more profoundly than one is rewarded for profits. Thus, loss aversion, identified by Kahneman and Tversky (1979), has a rational basis in the concavity of the curve of wealth.

[61] Problem adapted from Bernoulli (1954)

Prospect theory pathologized loss aversion, whereas, what Bernoulli is telling us is that the impact of losses is so punitive that it makes sense to overestimate the possibility of catastrophe, because losses can be existentially threatening. Thus, the inconsistencies between the predictions of expected utility theory and empirical evidence, Loss aversion is a rational adaptation to the prospect of catastrophic loss. This is even more clear where there is a chance of losing everything, in which case the geometric mean becomes zero.

A relatively recent investment example will show how the relationship between the disutility of losses and the utility of gains combines with the subjectivity of risk, in ways that can lead to confusion: in early May 2020, at the height of the Covid-19 pandemic, Warren Buffett's Berkshire Hathaway sold all its positions in Delta Air Lines, United Airlines, American Airlines and Southwest Airlines, for $9.3 billion, realising a loss of $5 billion. Between 2019 when Berkshire Hathaway took its positions in the airlines, and the decision to sell, American Airlines fell 62.9%, Delta Air Lines fell 58.7%, Southwest Airlines fell 45.8%, and United Airlines fell 69.7%. Financial services firm, New Constructs, called Buffett "short-sighted" on airlines, pointing to the industry's strong underlying economics and the obvious fact that airlines would fly again (Shuler 2020). A year later, American Airlines and Southwest Airlines were both up more than 80%, United Airlines and Delta Air Lines were up around 70%. An investor who bought at the May 25, 2020 bottom would have earned a 200% return from United Airlines, and a 190% return from American Airlines (Blikre 2021).

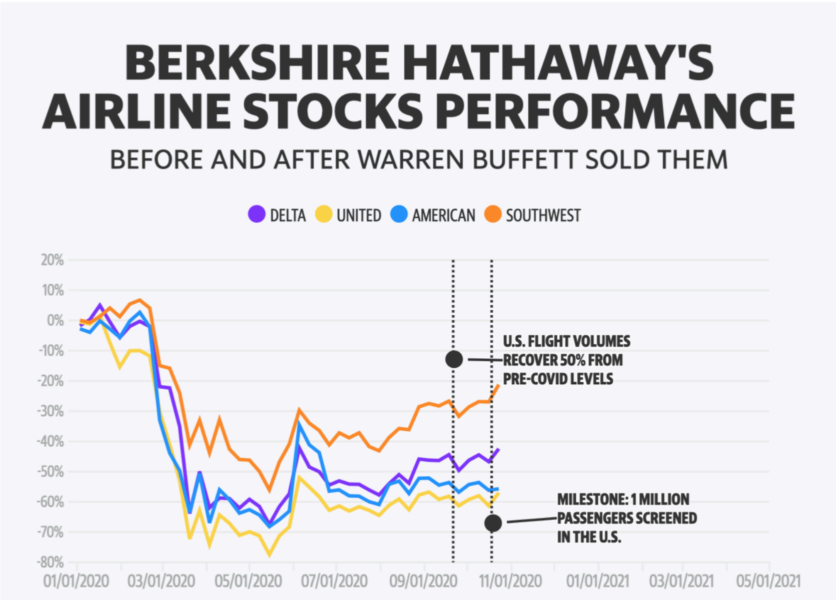

Figure 11: Berkshire Hathaway's Airline Stocks Performance. Source: Bilkre (2021)

That seems damning for a man considered by many as the greatest investor ever. However, simple arithmetic shows the subjectivity of risk: a loss of 70% requires a gain of 333% just to break even. The risk faced by holders of airline stocks was an opportunity for investors looking to buy, even if those investors looking to buy got their sums wrong and they did not get a 333% return. Risk does not exist as an objective category. Faced with catastrophic loss, both Buffett and the investors who bought into airlines, were right.

Bernoulli argues that "everyone who bets any part of his fortune, however small, on a mathematically fair game of chance is being "irrational". The impact of a loss may be so great as to be fatal. The disutility of loss is so high that one must only invest where one has a great advantage, or what value investors refer to as a "margin of safety". Fair games are actually deeply unfair and all investors face the possibility of gambler's ruin[62] even if they have a positive expected value. Loss aversion is intrinsic to estimations of risk. To not be loss averse is irrational.

Diversification and The Equity Premium Puzzle

Bernoulli showed that diversification can be derived from his theory and that diversifying increases the geometric mean wealth from exposure to a risky proposition. However, he also found that there are limits to the benefits of diversifications, as the costs of diversification rise

[62] The gambler's ruin problem refers to the eventuality of a gambler with a negative expected value going broke and was first posed and solved by Blaise Pascal in a 1656 letter to Pierre de Fermat (Edwards 1983), in which he said,

> Let two men play with three dice, the first player scoring a point whenever 11 is thrown, and the second whenever 14 is thrown. But instead of the points accumulating in the ordinary way, let a point be added to a player's score only if his opponent's score is nil, but otherwise let it be subtracted from his opponent's score. It is as if opposing points form pairs, and annihilate each other, so that the trailing player always has zero points. The winner is the first to reach twelve points; what are the relative chances of each player winning?

Huygens (1656) rephrased the problem as,

> Problem (2-1) Each player starts with 12 points, and a successful roll of the three dice for a player (getting an 11 for the first player or a 14 for the second) adds one to that player's score and subtracts one from the other player's score; the loser of the game is the first to reach zero points. What is the probability of victory for each player?

Johannes Hudde, Daniel Bernoulli and Pierre Rémond de Montmort all proposed solutions (Edwards 1983).

and, too much diversification can reduce gains. Diversification provides us with an interesting aside.

Rajnish Mehra and Edward C. Prescott (1985)[63] found that between 1889 and 1978, the average real return on equities was 6.98%[64] (with a standard deviation of 16.54%), compared to 0.8% for short-term, virtually risk-free bills, giving equities a 6.18% premium over bills (standard error of 1.76%) that cannot be explained by a reasonable level of risk aversion. These results violate the 1% equity premium that the standard general equilibrium model generates (Mehra and Prescott 1982; 1985). Adding to the intrigue, in the 1979 to 2005 period, Dimson, Marsh, and Staughton (2008) found that the equity premium had risen to 8.1%. Robert J. Shiller (1982) had earlier shown that asset return means and variances could only be explained by either a high coefficient of relative risk aversion or a large consumption variability out of step with empirical evidence. However, Francisco Azeredo (2014) found that a reasonable risk aversion under the Arrow-Debreu economy could produce a negative equity premium.

To date, no plausible explanation of the size of the equity premium has been found, despite numerous theories emerging, (Mehra and Prescott 2008; 2003; Cochrane 1998; Kocherlakota 1996; Mehra 2003), particularly those based on rare events (Rietz 1988; Julliard and Ghosh 2012), myopic loss aversion (Benartzi and Thaler 1995), equity characteristics (Palomino 1996; Holmström and Tirole 1998; Grant and Quiggin, 2004; Bansal and Coleman II 1996), tax distortions (McGrattan and Prescott 2001), implied volatility (Graham and Harvey 2007), information derivatives (Soklakov 2020; 2015), and market failure (Kocherlakota 1996).

Deconstructing the Equity Premium Puzzle

The puzzle was assessed by comparing the returns of the S&P 500 with those of treasury bills[65]. This results in a false equivalency: one is a portfolio with clear selection criteria and a large number of investments, whereas the other is a single instrument. Secondly, the S&P 500 can be traced back to 1923 when the Standard Statistics Company developed an index of 223 U.S. companies. The S&P 500 itself was only created on March 4, 1957. So for the

[63] An earlier version of this paper, (Mehra and Prescott 1982), obtained similar results.

[64] The size of the equity premium depends on the time period studied, but over the long-term, it is around 5% to 8%.

[65] In the 1931-1978 period, 90-day government treasury bills were used; in the 1920 to 1930 period, treasury certificates were used; and 60 to 90-day prime commercial paper was used for the pre-1920 period.

first 68 years of the study period, the hypothetical investor could not have invested in the S&P 500. Indeed, it was only on August 31, 1976, when the Vanguard Group offered retail investors the first mutual fund that tracked the S&P 500, could retail investors invest in the S&P 500. On a practical basis, the puzzle is illusory because the hypothetical investor of all but the last year or so of that period, could not invest in the S&P 500.

Now, if we take the puzzle on its own terms, the size of the equity premium is, of course, smaller. We can quickly see this by converting the arithmetic returns to geometric returns, using Booth and Fama (1992)'s derivation of the relationship between the geometric return, g, and the arithmetic average return, r, such that,

$$g = r - \tfrac{1}{2}\sigma^2 \tag{24}$$

we can calculate the geometric return of the S&P 500 for that period as,

$$g = 6.98\% - \tfrac{1}{2}*16.54\%^2 = 5.61\% \tag{25}$$

As we have said, the puzzle rests upon our accepting the comparison of a portfolio with a single instrument. If, however, we compare the typical stock with Treasury bills, the equity premium puzzle melts away. Hendrik Bessembinder (2018) found that four out of seven stocks in the CRSP database, had lifetime buy-and-hold returns lower than one-month Treasuries, between 1926 and 2018. Furthermore, just 4% of publicly traded companies in the database account for the net gain for the U.S. stock market in that same period, with the other stocks having returns equal to those of Treasury bills. More precisely, 96% of stocks, if held across their lifetime, cannot match the returns of Treasury bills. While Bessembinder does not say this, we can conclude that the equity premium puzzle is thus non-existent. What the equity premium puzzle does show is that a portfolio of assets with average return *x*, can achieve higher returns than a single asset with average returns *x*.

Caius the Merchant Discovers that Insurance Increases Wealth

The insurance puzzle describes the strange coincidence of supply and demand for insurance contracts in an expected value world. An insurer is rationally obliged to maximise their net premium, which, in order for a contract to be signed, implies that an insurance buyer must act irrationally in terms of expected value, and purchase insurance at prices against their interests. The insurance paradox, which gained renown with Rothschild and

Stiglitz (1976) work, highlights the inability of economics to explain the existence of a competitive market for insurance contracts.

A person with non-risky W wealth, facing the possibility of an adverse event, seeks to insure a good of value *v*, by paying an insurance premium, βx, where $0 < \beta \leq 1$, in return for which he will gain coverage, *x*. If the adverse event occurs, this person will have wealth, *W*, and if it does not occur, wealth *W+v*. They face two possible states, "no adverse event", "adverse event", (W + v, W), without insurance, and which, with insurance, is transformed to $(W + v - \beta x, W + x - \beta x)$. The probability of the accident, p, is weighed against the probability that the accident won't occur, (1-p). The estimated wealth, E(W) of the insured party can be calculated as a linear function on x,

$$E(W) = (1-p)(W + v - \beta x) + p(W + x - \beta x) = w + x(p - \beta) + (1 - p)v \tag{26}$$

And the expected profit of the insurer can be defined as,

$$E(\pi) = (1-p)\alpha_1 + p\alpha_2 = \alpha_1 - p(\alpha_1+\alpha_2) = \alpha_1 - p\bar{\alpha}_2 \tag{27}$$

which leads us to,

$$E(W) = W + p(\bar{\alpha}_2 - pd) - \alpha_1 \tag{28}$$

Bernoulli's (1954) neglected work provides an elegant and powerful answer to the insurance paradox: insurance contracts increase wealth. Bernoulli provides a simple analogy: Caius is a merchant who has bought a consignment of goods in Amsterdam which he plans on shipping to Saint Petersburg, to earn 10,000 rubles upon delivery, however, 5% of ships that hazard this journey are lost every year. The lowest cost insurance he can get is priced at 800 rubles, which is excessively priced,

The expected value of that contract is,

$$\begin{aligned} EV_I &= (5\%*9200)+(95\%*-800) \\ &= -300 \text{ rubles.} \end{aligned} \tag{29}$$

He is rationally obliged to walk away from this offer. However, given the framework Bernoulli has provided, what Caius actually has to do is to determine at what level of wealth he can abstain from purchasing insurance.

Given w_0 as his prior fortune, his wealth in the event of safe travel is given as,

$$W_s = (w_0 + 10{,}000)^{95} * w_0^{5})^{1/100} = (w_0 + 10{,}000)^{19} w_0)^{1/20} \quad (30)$$

Insurance guarantees him a certain fortune, CF, w_0 + 9,200, which can be represented by

$$CF = w_0 + 9{,}200 \quad (31)$$

From this, Bernoulli surmises that

$$(w_0 + 10{,}000)^{19} w_0 = (w_0 + 9{,}200)^{20} \quad (32)$$

From this, it is seen that,

$$w_0 = 5{,}043 \text{ rubles} \quad (33)$$

With a prior wealth of 5,043 rubles and above, Caius can afford to remain uninsured. With a lower prior wealth, insurance is necessary. The insurer, on the other hand, with a prior fortune of w^i_0, expressible as,

$$((w^i_0+800)^{19}(w^i_0-9200))^{1/20} = w^i_0 \quad (34)$$

From which w^i_0 = 14,243.

Bernoulli demonstrates that, ceteris paribus, by buying insurance, an investor may increase their W_B, while at the same time, an insurer improves theirs[66].

Bernoulli's Error

Bernoulli's paper is not without faults. He makes one galling error. Bernoulli's theory begins with the assumption that no fee is paid to participate in this risky proposition. However, when he tackles the question of how much one must be prepared to pay to enter this gamble, he, as Peters (2020) has noted, makes a logical error: given a maximum price R_m, which must

[66] Ole Peters (2017) provides a formal proof of Bernoulli's argument that insurance grows geometric wealth.

be deducted from one's prior wealth, to give us post-fee wealth, W_F, which can be defined as,

$$w_0 - R_m + \pi_i = W_F \tag{35}$$

This of course cannot be reduced to wo + π_i. Given a fee, his curve shifts leftward and of course,

$$W_F < w_o + \pi_i \tag{36}$$

for the same π_i, except where $R_m = 0$. Bernoulli's fee suggests that it is impossible to make a loss from the bet. So, while Bernoulli speaks to the nature of risk, and the importance of the geometric mean, it was left to his heirs, Shannon (1971), Kelly, Jr. (1956) and Thorp (1969; 1966; 1971; 1997), to deepen his theory and answer the question of how investors should invest with the prospect of ruin.

Bernoulli's Heirs and Investing Under the Prospect of Ruin

Chaude Shannon and Choice, Uncertainty and Entropy[67]

Claude Shannon's (1971) information[68] theory provides a starting point to a question that Bernoulli left unanswered, or answered very badly: "How should an investor invest in a world of almost certain ruin?". Shannon's seminal paper, *The Mathematical Theory of Communication*, treated the "fundamental problem of communication", where communication is defined as all procedures by which one tries to influence another person, from speech to ballet (Weaver 1971). This fundamental problem, or what Warren Weaver (1971) refers to as a "Level A" problem, is that of being able to exactly or approximately identify a message from an information source, based on the received signal emanating from the communication channel[69]. So, for instance, if I send a text message with the words "Grand Inquisitor", the

[67] Title borrowed §6 of Shannon (1971).

[68] Information is used in a special sense to mean "a measure of one's freedom of choice when one selects a message" (Weaver 1971, p. 4).

[69] Essentially, an information source selects a message out of a set of possible messages, and this message is then changed and sometimes encoded into a signal by a transmitter, over some communication channel, with a capacity measured in bandwidth. Some noise source corrupts the

recipient of this message should be able to receive that exact message, or something close to it, for instance, “Gr?nd Inquisto”. A message does not have to have any meaning, and therefore, the concept of information is itself divorced from meaning. So, the fundamental problem of communication holds if all I type is, “gvjcddshjg”. The question is one of accuracy, and it is for this reason that Shannon notes that the “semantic aspects of communication are irrelevant to the engineering problem”[70]. Nevertheless, the engineering aspect of communication do have great implications for the semantic problem of the accuracy of symbols in conveying meaning, the semantic or Level B problem, and the problem of the effectiveness of the meaning in influencing behaviour in the desired manner, the effectiveness or Level C problem[71] (Weaver 1971).

The fundamental problem of communication compels any system to be able to operate across a set of possible messages, not simply the message that has been sent. This is because, prior to its use, it is not known what message will be transmitted through this system. Properly speaking, then, information refers to the degree of freedom one has in choosing a message. So, for example, if one were writing a message to a friend named David, and began that message with “Hello”, the degree of freedom I have in naming him is limited to variations of David, for example, “Hello David”, or “Hello Dave”, to variations of things that I call David, for example, “Hello Old Sailor!”. Similarly, “Hello David” has variations such as, “Hi Dave”. A system is agnostic about the selection I will make and has to be able to cope with all these variations.

The amount of information is measured by the logarithm of the number of available choices, specifically, logarithms to base 2, so that a 2-choice situation yields unity, or what J.W. Tukey termed a “bit” from “binary digit” (Shannon 1971; Weaver 1971). So, if a person has 16 ($=2^4$) different messages from which they are equally free to choose, then they have 4 bits of information to choose from[72] Weaver 1971). The number of bits of information is, therefore, defined as,

$$\log_2 2^N = N \tag{37}$$

signal, which creates difficulties for the receiver, who has to convert the signal back into a message, before handing it over to its intended destination.

[70] Shannon 1971, 1

[71] The effectiveness problem was treated by von Neumann and Morgenstern (1953)’s development of game theory, and, in the Soviet Union, by Vladimir Lefebvre’s theory of reflexive control (Thomas 2004).

[72] $16 = 2^4 \rightarrow \log_2 16 = 4$

The choice of message available to the information source is governed by probabilities that are themselves dependent on prior choices. So, for instances, within the context of a discussion about one's favourite sausages, if one said, "I love boerewors", it would be absurd for the recipient, despite any noise corrupting the message, from concluding that one meant to say, "I love boreholes". The probabilities would narrow the range of choices to a set of sausages, and even if the recipient had never heard of boerewors, they could still conclude with a high degree of confidence that "boerewors" are a type of sausage. Indeed, if the recipient only received the first two words, "I love", and the third was corrupted, the probabilistic structure of the message would narrow down the possible options. In other words, communication is a Markov process[73]. And the most important of these Markov processes are *ergodic processes*.

Shannon, after a suggestion from John von Neumann referred to the amount of information (or freedom of choice) produced by this Markov process as entropy (McIrvine and Tribus 1971)[74]. Shannon's entropy is directly analogous to entropy in thermodynamics[75], such that, given a set of possible events with probabilities of occurrence, (p_1, p_2, …., p_n), H(p_1, p_2, …., p_n) is a measure of the available freedom of choice, or the uncertainty of the outcome, such that the formula for Shannon's entropy is,

$$H^{76} = -K \sum_{i=1}^{n} p_i \log p_i, \qquad (38)$$

---

[73] A Markov process or Markov chain is named after Andrey Markov and refers to a type of stochastic process wherein each event or in this instance, symbol, in a sequence of events has a probability dependent on the state achieved in the prior event.

[74] Shannon said, "I thought of calling it "'information', but the word was overly used, so I decided to call it "uncertainty". [...] Von Neumann told me, 'You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, nobody knows what entropy really is, so in a debate you will always have the advantage.'"

[75] Entropy was first calculated by Rudolf Clausius in 1865 (Weaver 1953) and is associated with H-theorem Ludwig Boltzmann's development of the H-theorem in 1872 (Pathria and Beale 2022; Weaver 1971; Shannon 1971). The second law of thermodynamics states that entropy is always increasing.

[76] This formula is motivated by its ability to meet the following criteria:

where Shannon's entropy takes the same symbol, H as entropy in thermodynamics, and K[77] is a constant. H is a measure of the randomness of a situation and overtime, H increases. The base of the logarithm changes according to the application. Where base 2 gives bits, base e gives natural units, and base 10, bits, "bans" or "hartleys".

A simple instance will illustrate how H works. Shannon shows that given two outcomes, with probabilities p, and q = 1 - p,

$$H = -(p \log p + q \log q), \qquad (39)$$

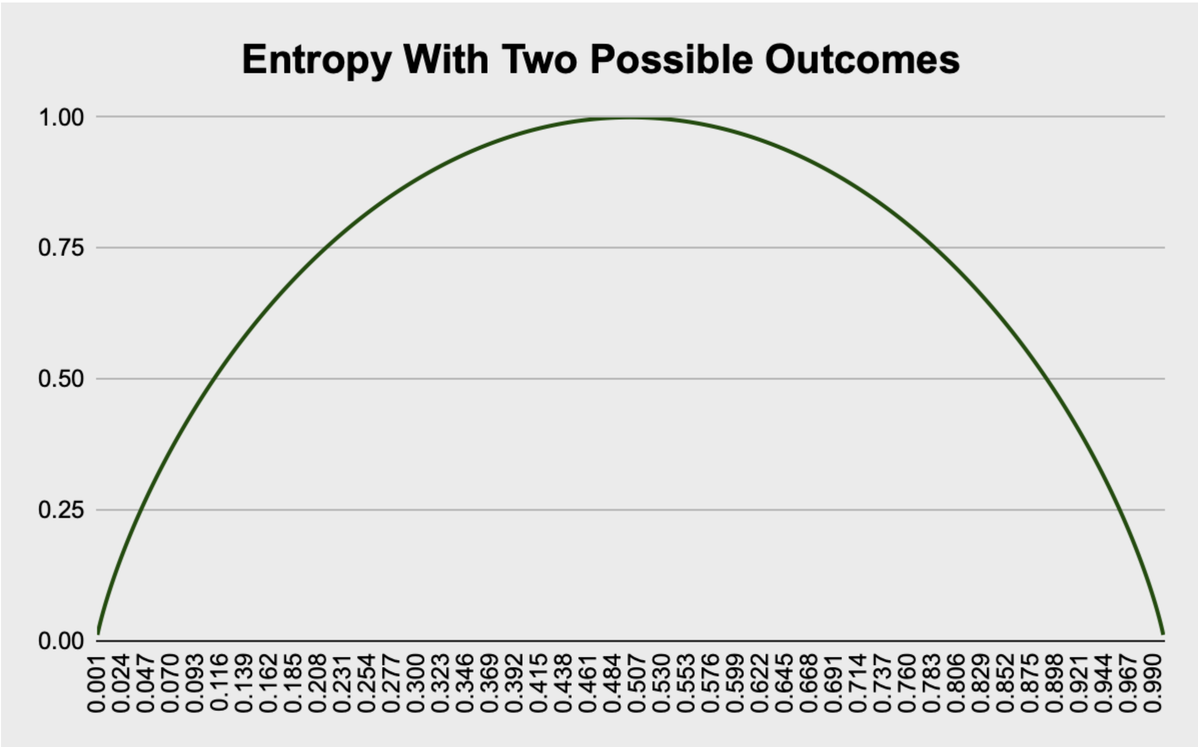

Figure 12: Entropy With Two Possible Outcomes

1. H is a continuous function of p_i. This is so that an arbitrarily small change in the probability distribution does not result in a large change in uncertainty (Jaynes 2003).
2. H should correspond to common sense notions that, if all p_i are equal, with a rise in the number of possible outcomes, uncertainty increases, i.e. H is a monotonic increasing function.
3. H is consistent such that its value is the same regardless of how it is calculated.

[77] K represents some unit of measure.

As the chart shows, H is at its highest when p = q, and a person can freely choose between the two options. This result is extendable to the rule that H is at its greatest when the p_i are all equal or approximately so, giving one the maximum of freedom of choice (Shannon 1971; Weaver 1971). Where any one choice has a probability that is larger than the others, the greater this probability, the more choice is narrowed toward it. H can also be increased by raising the number of choices, assuming that the choices have an equal or nearly equal probability.

The capacity of a communication channel is, then, related to the amount of information that it conveys. In other words, the better a communication channel is at transmitting whatever comes out of an information source, the greater its capacity. The capacity of a communication channel is measured in terms of the number of symbols transmitted per second, n, and the bits of information, s, that are freely chosen from 2^S symbols.

Information is, then, what one doesn't know, it's about unpredictability. Weaver (1971) makes it clear just how common sense H is, when he says that one could say of an information source, "This situation is highly organised, it is not characterised by a large degree of randomness or of choice - that is to say, the information (or the entropy) is low". Information is what is unknown and unpredictable.

Having measured entropy, one can then find the ratio between actual and maximum entropy, while still using the same symbols. Relative entropy is a measure of how much an information source can be compressed if encoded using the same alphabet. One minus relative entropy is the redundancy of the information source. Redundancy is that part of an information source that is determined by the statistical structure of that source, with the rest determined by free choice. Shannon calculated, for instance, that English had a redundancy of 50% up to eight letters, so that half of what one writes or says in English is determined by the statistical structure of the language and the other half by free choice.

Shannon realised that the statistical nature of the information source was important in solving the fundamental problem of communication, because it is impossible to create a communication system that will perfectly process every single message it receives, but it should be able to handle every conceivable type of message, especially the most likely types, with efficiency declining for the most rare message types (Shannon 1971; Weaver 1971). Thus, the theory of information deals with Level A analysis. So, the capacity of a communication channel is more appropriately viewed with respect to the information transmitted and not to the number of symbols. In this channel, a transmitter accepts messages turning it into a signal, encoding it, passing it over the channel, before it is decoded by the receiver.

The above leads to the fundamental theorem of a noiseless channel: supposing an entropy source with entropy H bits per symbol, and a channel with capacity C bits per second, the source's output can be encoded such that the channel can transmit symbols at an average rate of $C/H - \epsilon$, where ϵ is arbitrarily small (Shannon 1971; Weaver 1971). The transmission rate can never exceed C/H. As we have seen, the statistical character of the symbols that the channel actually transmits is determined by what one intends on transmitting and the channel's capacity.

In the absence of complete signal freedom, signal entropy can be achieved that is larger than other statistical signal structures. In this scenario, the signal entropy is equal to the channel capacity. The best transmitter will code a message such that the signal possesses the optimum statistical characteristics for the communication channel it is transmitting over, thus maximising signal or channel entropy and equating it to the channel capacity, C. In achieving the maximum transmission rate, C/H, one is forced to accept increasingly longer delays in the coding process.

Given that information represents the degree of freedom one has in selecting a message, the more freedom one has, the greater the information and the greater the uncertainty about the selected message. Noise affects information by distorting information, introducing errors, and unnecessary material, increasing the degree of uncertainty. The greater the uncertainty, the greater the information. In this sense, there is undesirable uncertainty brought about by the presence of noise, whereas when a sender enjoys freedom of choice, that gives the sender desirable uncertainty (Weaver 1971).

Now, returning to our earlier discussion of relative entropy, we can say the following. Given two sets of symbols, such as those transmitted from the information source and those received, we can see that their probabilities are interrelated. In a noiseless communication channel, the received signals would be the same as the transmitted message symbols. In a noisy communication channel, the probabilities of those received symbols would be tilted toward those that correspond to the transmitted message symbols. Although we might know the amount of noise in the communication channel, this knowledge is useless to the recipient because that knowledge does not also say where the noise is connected. Thus, the recipient is ignorant of the rate of transmission of information within a noisy communication channel. If the noise is extreme enough, the received symbols will be completely independent of the transmitted symbols, and a communication system based on chance might be more accurate than the communication system at hand.

Clearly, the amount of information transmitted could be corrected by the amount of missing information in the received signal. Shannon used conditional entropy of the message as a measure of the missing information. This allows us to calculate the rate of actual transmission of information, R, by subtracting the average rate of conditional entropy, from the entropy of the source, such that,

$$R = H(x) — H_y(x) \tag{40}$$

where H(x) is the entropy of the message source, and $H_y(x)$ is what Shannon called the conditional entropy, or, equivocation, which is a measure of the average ambiguity or uncertainty of the received signal. The average uncertainty in the message source when the received signal is known, is the same as the average uncertainty of the received signal when the message sent is known. In the absence of uncertainty, the latter is zero.

$$H(x) — H_y(x) = H(y) — H_x(y) \tag{41}$$

where H(y) is the entropy of received signals and $H_x(y)$ is the uncertainty of the received signals if the transmitted messages are known, or the undesirable part of the received information, such that the right-hand side of the equation represents the useful information transmitted through the noisy communication channel.

Thus, the capacity, C, of a noisy communication channel, is equal to the maximum rate at which it can transmit useful information. The transmission rate can be maximised with the appropriate coding. Regardless of how well we encode the signal, there is always undesirable uncertainty, or equivacation, about the actual transmitted message, and this equivocation is always greater or equal to H(x) - C. At least one code can reduce the undesirable uncertainty to an amount that is arbitrarily greater than H(x) - C. Nevertheless, the minimum undesirable uncertainty has a limit, regardless of how the coding process can be improved. This is the limit of the dependability of a noisy communication channel.

John Larry Kelly, Jr. and the Kelly Criterion

John Larry Kelly, Jr. became acquainted with Shannon when the two worked at Bell Labs (Poundstone 2005). He found an innovative application to Shannon's information theory. Kelly, Jr. (1956) posed a thought experiment: suppose a gambler is given advance notice of the outcome of a risky proposition, allowing him to gamble at the original odds, before the new information has been priced in. Although he trusts these messages, they are not without risk, even though they give him an edge.

Kelly proposed that in a *noiseless* binary channel, equivalent, one must note, to Knightian risk, this fortunate gambler could, for instance, gamble on the outcome of a series of baseball games between evenly matched teams, getting even money bets despite knowing the outcome of the result. His more logical gambling strategy would be to place as much money as he could, on the certain outcome he knew, his capital growing exponentially by 2^N times, where N represents the number of bets placed. So, if our fortunate gambler received weekly messages it would be equivalent to an investment with a return of 100% interest every week. Kelly denoted the exponential rate of growth of the gambler's capital by *G*, where,

$$G = \lim_{N \to \infty} 1/N \; log V_N/V_0 \tag{42}$$

Where V_N is the gambler's capital after N bets, V_0 his starting capital, and the logarithm is to base 2.

Supposing, however, that rather than a noiseless binary channel, one in fact was working with a noisy binary channel, in which each transmitted symbol has a probability, *p*, or error, and *q* of correct transmission. Rephrased, Kelly asked himself what would happen if a gambler had an uncertain edge. If this gambler decided to maximise the expected value of his capital, $E[V_N]$, denoted by,

$$E[V_N] = (2q)^N V_0 \tag{43}$$

our fortunate gambler would find that if he bet everything, he would face gambler's ruin the moment he faced an adverse result. On the other hand, Feller (1966) showed that, self-evidently, a policy of minimising the possibility of ruin by placing the minimum possible bet, results in minimising the expected gain. The rational thing to do, instead, is to bet a portion, *l*, of his capital, such that,

$V_N/ = (1 + l)^W(1 - l)^L V_0$ (44)

where W is the number of wins, and L the number of losses, in the bet, so that,

$$G = \lim_{N \to \infty} [W/N * log(1 + l) + L/N * (1 - l)]$$

$$= q*\log(1 + l) + p*\log(1 - l) \text{ with probability one} \quad (45)$$

Kelly showed that maximising G with respect to *l*, leads to,

$$G_{max} = 1 + p*\log p + q*\log q$$

$$= R \quad (46)$$

Which is of course (41). The gambler's rational course of action is to maximise the logarithm of their capital. More profoundly, Kelly's finding is that the maximum rate of return is equal to the flow of inside information, and a few bits of information can generate outsized returns. Kelly showed that this result is extendable to situations where there are multiple input symbols with different probabilities, such that

$$G_{max} = H(X) - H_x(y) \quad (47)$$

In other words, the gambler would be able to compound their wealth at an exponential rate, without risking ruin.

Where there is a "track take", that is, where the track sums up the winning wagers for a race, and deducts fees for expenses, distributing the rest to the winners, Kelly showed that the gain depends on the size of the bets placed on the winning horse. When there is no track take, if, for example, a tenth of the bets wagered go to the winning horse, Stormy Weather, in other words, Stormy Weather is paying 9:1, then the winning gamblers will increase their capital by 10 times. A gambler with inside information can bet all their capital at each race, using a subjective estimate of each horse's odds of winning to apportion your capital. Thus, certainty is achieved in that one horse will win, guaranteeing some return. In doing so, the gambler ignores the posted odds, favouring instead his superior odds based on inside information. Suppose, for example, that Stormy Weather is, according to inside information, guaranteed to win. The rationally obligatory thing to do is to place everything on that certain winner, and get paid according to the posted odds. Edge implies concentrated betting to maximise returns.

In the real world, Shannon's communication system and Kelly's betting system work through uncertainty. Kelly proved that by factoring in equivocation, a gambler should place bets consistent with their probabilities. This policy, given an edge, results in the maximisation of the gambler's logarithm of capital.

Kelly's gambling system results in concentrated bets. The implication is that, absent an edge, a person should, as Bernoulli advised, diversify, but given an edge, a person should concentrate their bets based on their beliefs. Kelly, essentially, advises gambler's, given the possibility of losing everything, that they bet a fraction, f^*, of their capital, such that,

$$f^* = p - q/b = p - (1 - p)/b \tag{48}$$

where p is the probability of a win, and q is the probability of a loss, and b is the proportion of the bet that is gained from a win, i.e.

If only part of one's capital is threatened, then, the appropriate formula for investors is,

$$f^* = p/a - q/b \tag{49}$$

where *a* is the proportion of the investment lost in the event of a negative outcome.

Now, the Kelly criterion, as its known, is true for known outcomes, but it has proved a useful investment tool for some of the world's greatest investors, such as Warren Buffett, Bill Gross, and Jim Simons (Zuckerman 2019; Benello, van Biema, and Carlisle 2016; Thorp 1966; 1997). It also echoes Bernoulli's finding that, given a choice of investments, one should select those with the highest geometric mean. Investors may of course face constraints that prevent them from investing using the geometric mean as their North Star, (Poundstone 2005), and even users of the Kelly criterion typically invest less than the Kelly criterion advises, so-called fractional Kelly betting, in order to reduce volatility and account for errors in their estimation of their edge, but the criterion does speak to the nature of risk. In the face of the "concavity of the curve", investors must scale back when they incur losses, and double down when they win, because the disutility of a loss is more impactful than the utility of a gain.

Bessembinder's (2018) surprising results about stock returns *qua* Treasury bills, can be overcome in light of Kelly: the longer the typical stock is held, the greater the chance that its geometric returns will collapse toward those of Treasury bills. Consequently, investors have to rebalance as frequently as costs allow, in order to maximise their geometric returns. Shannon showed that it was even possible to earn a positive geometric return from purely random markets, through rebalancing[78] (Poundstone 2005). The investor who holds the typical stock, that is, 96% of U.S. stocks, across their lifetime, will not earn a rate of return greater than Treasury bills. Elwin Berlekamp, who studied under Shannon and worked with Kelly, argued that not only is an investor obliged to rebalance frequently, to avoid the rising risk of loss, but doing so minimised the consequences of each trade, reducing portfolio risk (Zuckerman 2019). He likened the ideal fund (he developed the Medallion Fund's[79] investment philosophy) to a casino that handles so many bets that it only needed a sliver of profit from a majority of those bets, to be massively profitable, bringing the law of large numbers to their side. Importantly, in speaking about becoming a casino, he was referring to bets across time, rather than in parallel. The secret to overcoming the near-certainty of being unable to beat the market, is bet sizing and diversification across time.

(48) and (49) imply that if an investor or gambler does not have any inside information, they should not wager anything, one's edge is zero, or, more appropriately, given track take, or investment fees, it is negative. Centuries after Bernoulli, Kelly's criterion warns us about the perils of investing.

Without discussing ergodicity, Kelly demonstrates the impact of time on investments, saying,

> The gambler introduced here follows an essentially different criterion from the classical gambler. At every bet he maximizes the expected value of the logarithm of his capital. The reason has nothing to do with the value function which he attached to his money, but merely with the fact that it is the logarithm which is additive in repeated bets and to which the law of large numbers applies. Suppose the situation were different; for example, suppose the gambler's wife allowed him to bet one dollar

[78] This experiment has been called, "Shannon's demon".

[79] Jim Simons' Medallion fund, now Renaissance Technologies, earned 66% annualised returns (before fees) and 39% annualised returns (after fees) between 1988 and 2018. $1 invested in the fund in 1988 would be worth more than $20,000 (after fees) in 2018, compared to $20 for the S&P 500, and $100 for Warren Buffett's Berkshire Hathaway. Renaissance Technology levied a 5% management fee and 44% performance fees.

> each week but not to reinvest his winnings. He should then maximize his expectation (expected value of capital) on each bet. He would bet all his available capital (one dollar) on the event yielding the highest expectation. With probability one he would get ahead of anyone dividing his money differently.

In other words, if one's capital is reinvested in the market, that is, it flows through time, then one should maximise the logarithm of capital, but in the absence of time, when winnings are taken out of the market, the appropriate way to invest is by maximising, not geometric returns, but arithmetic returns. In fact, Edward O. Thorp (1969), believed that "the Kelly criterion should replace the Markowitz criterion", showing that Markowitz' mean-variance criterion was partially incompatible with the Kelly criterion. Once we introduce time into the world of gambling and investing, by reinvesting profits, we are confronted with gambler's ruin, and, conversely, the opportunity to take advantage of the multiplicative nature of wealth.

Although the Kelly criterion is widely used in investing and gambling communities (Benello, van Biema, and Carlisle 2016; Thorp 1966; 1997), it has not been met with much welcome in the academic economics. Paul Samuelson's (1969) dismissal of the Kelly criterion, as faulty and imprecise, is reflective of economist's response to the Kelly criterion. Henry Latané[80] and Nils Hakansson favoured the Kelly criterion, while Mark Rubinstein adopted it and then recanted his position (Poundstone 2005). Henry Markowitz revised his 1952 theory of portfolio selection in light of Latané's work on portfolio selection using the geometric mean, but his revision remains the least read part of his book on *Portfolio Selection*. Samuelson (1971, 2494) referred to the Kelly criterion as a fallacy, but in doing so, assumed its central point that wealth is maximised at the geometric mean (calling it an "indisputable truth", saying,

> Theorem. If one acts to maximize the geometric mean at every step, if the period is "sufficiently long," "almost certainly" higher terminal wealth and terminal utility will result than from any other decision rule.
>
> …. From this indisputable fact, it is tempting to believe in the truth of the following false corollary:
>
> False Corollary. If maximizing the geometric mean almost certainly leads to a better outcome, then the expected utility of its outcomes exceeds that of any other rule, provided T is sufficiently large.

[80] Latané (1959) treats the problem of portfolio selection with a geometric mean criterion.

In other words, the Kelly criterion is the best criterion for maximising terminal wealth, but it does so by making tradeoffs which not every investor may be prepared to make. Specifically, the Kelly criterion is extremely loss averse, shunning every chance of extermination and seeking out long-term performance, even at the risk of short-run sub-par performance. This, again, goes against orthodoxy, which holds that high returns are a function of taking on more risk. It also narrows our vision, or more precisely, makes it clearer, seeing that risk is, in the realm of investing, not about missing a target, it's about the possibility of a permanent loss of capital. Furthermore, as Breiman (1975) showed, the geometric mean criterion is the fastest way to grow terminal wealth.

Latané (1959) referred to the policy of portfolio selection using a geometric mean criterion as myopic, "near-sighted", because it does not attempt to peer into the future, instead, focusing primarily on making decisions based on near-term probability distributions, rebalancing as those distributions materially change. It is Latané's work that influenced Markowitz to embrace the geometric mean criterion, after becoming aware of Latané's research in 1955-56 (Poundstone 2005). Markowiotz immediately realised that the mean-variance theory assumed a static, single investment period, whereas, ergodicity theory tells us that that is not part of the nature of risk.

Conclusion

The Knightian consensus established the view that risk is defined by two features. The first feature is stated explicitly: risk is measurable, and knowable, and, Knight implies, knowable to a degree that eschews an error rate. The second feature of risk, according to this consensus, is implied, and it is that risk is a forecasting problem. However, the question is, "What are we forecasting?" Knight's view sees the range of all possible outcomes as being at issue. This thesis has tackled those two features of risk.

It has been argued in this thesis that the subjective and the logical interpretations of probability, along with the lessons of the pre-Pascalian consensus, compel us to view risks as being subjective, epistemic categories, rather than as objective categories that leave us with the simple task of merely measuring them. As such, we cannot peel uncertainty away from risk. This is especially true because our estimations of risk involve value judgements that are imprecise, difficult to disentangle, and which may even be arbitrary. Investors and managers must also deal with the problem of weighing evidence, a problem which is often not resolvable with recourse to some number. Indeed, even if we assume that risk is objective, risk models and theories of risk are epistemic categories, and they create uncertainty around our interpretation of risk. Uncertainty cannot be disentangled from risk and not all risks can be measured. There is an argument that we should assign numbers to describe our degree of belief in a proposition, in order to be able to clearly communicate this degree of belief, but this number is very far from the kind of objective probabilities that the Knightian consensus advances.

Bernoulli (1954) deepens our sense of subjectivity, by leading us to the view that risk is not just subjective from an epistemic point of view, but also in terms of prior wealth and the position an investor is taking. Thus, risk cannot be calculated as an universal, objective objective. Risk for one party is not necessarily risk for another party. Indeed, exchange between parties can only exist if risk is subjective. Borrowing from de Finetti (2017), we can say that risk does not exist, by which one means that risk is not an objective category.

We can speak of a modified Knightian continuum of probability situations that are either measurable or to which a number can be assigned, even if for purely communicative purposes, and, of another aspect of risk, which deals with arguments, and the weighing of evidence, for which it is impossible to assign a number. Developing a theory of probability is, however, outside the scope of this treatise. Nevertheless, sufficient grounds have been given for embracing a broader view of probability theory, so that risk managers and investors can better arm themselves against “the slings and arrows of outrageous fortune”[81].

Bernoulli’s deepest insight was to show that investors face a concave curve of wealth, and that is the objective part of risk: the multiplicative nature of wealth. Given the multiplicative nature of wealth, the disutility of losses is greater than the utility of gains. It, therefore, makes sense to be loss averse. Indeed, this is rationally obligatory given the nature of risk. It is not a bias, or a sign of deficiency. Losses are hard to come back from. This principle is so important that Bernoulli referred to it as “nature’s admonishment to avoid the dice”. Losses are so punitive that one is faced with gambler’s ruin. In answer to that, one must diversify and insure one’s wealth, as a way of climbing quickly up the curve of the logarithm of wealth. Although flawed, Bernoulli’s work provides insights into the equity premium puzzle, the insurance paradox, and, of course, the Saint Petersburg paradox. Bernoulli did not provide a useful guide to invest in this perilous world as an active investor, but Claude Shannon and John L. Kelly, Jr. did.

[81] Hamlet, Act III, Scene 1, William Shakespeare (2003).

Shannon's communication theory, and his solution to the fundamental problem of communication, served as the basis of the Kelly criterion, which echoes many of Bernoulli's ideas. The Kelly criterion confirms Bernoulli's finding that the best way to grow wealth is by maximising the geometric mean. Kelly also showed that given the punitive nature of losses, if an investor does not have any edge, it makes no sense to invest, because even if that investor earned the market return, that return would, practically, be lower than the market return, given fees. The Kelly criterion is extremely loss averse, and yet, it proves Bernoulli's other finding: that returns are greatest when risks are reduced.

Although the Kelly criterion has become popular among investors, it still does not have much standing in academic finance. This is a shame. It is also a shame that, despite anticipating Kelly and Thorp by several centuries, Bernoulli's contribution to the discipline of investing has been forgotten.

We can conclude by advancing the view that risk should be viewed myopically, as being, primarily, the possibility of loss, and, secondarily, the possibility of missing some target rate of return. Probability, the mathematical and logical models that we develop to assess this risk, invites us to see risk as being subjective, containing both quantitative and qualitative aspects, and defined primarily by loss, and secondarily by the possibility of missing a target rate of return. Wealth, being multiplicative, leads to losses having a greater impact on portfolios than gains, making it rationally obligatory to be loss averse. Loss aversion is not simply a matter of avoiding losses, it is the basis of any rational strategy to build wealth.

Appendix

Appendix 1: Data for the Capped Saint Petersburg Wager

The following is a table of the data used to generate "Figure 8: Capped Saint Petersburg Wager":

Number of Tosses	Probability	Payoff	Expected Payoffs	Expected Value
1	0.5	$ 1.00	0.5	0.5
2	0.25	$ 2.00	0.5	1
3	0.125	$ 4.00	0.5	1.5
4	0.0625	$ 8.00	0.5	2
5	0.03125	$ 16.00	0.5	2.5
6	0.015625	$ 32.00	0.5	3
7	0.0078125	$ 64.00	0.5	3.5
8	0.00390625	$ 128.00	0.5	4
9	0.001953125	$ 256.00	0.5	4.5
10	0.0009765625	$ 512.00	0.5	5
11	0.00048828125	$ 1,024.00	0.5	5.5
12	0.000244140625	$ 2,048.00	0.5	6
13	0.0001220703125	$ 4,096.00	0.5	6.5
14	0.00006103515625	$ 8,192.00	0.5	7
15	0.00003051757813	$ 16,384.00	0.5	7.5
16	0.00001525878906	$ 32,768.00	0.5	8
17	0.000007629394531	$ 65,536.00	0.5	8.5
18	0.00000381469	$ 131,072.00	0.5	9

	7266			
19	0.000001907348633	$ 262,144.00	0.5	9.5
20	0.0000009536743164	$ 524,288.00	0.5	10
21	0.0000004768371582	$ 1,048,576.00	0.5	10.5
22	0.0000002384185791	$ 2,097,152.00	0.5	11
23	0.0000001192092896	$ 4,194,304.00	0.5	11.5
24	0.00000005960464478	$ 8,388,608.00	0.5	12
25	0.00000002980232239	$ 16,777,216.00	0.5	12.5
26	0.00000001490116119	$ 33,554,432.00	0.5	13
27	0.000000007450580597	$ 67,108,864.00	0.5	13.5
28	0.000000003725290298	$ 134,217,728.00	0.5	14
29	0.000000001862645149	$ 268,435,456.00	0.5	14.5
30	0.0000000009313225746	$ 536,870,912.00	0.5	15

Bibliography

Alchian, Armen A. ‘Uncertainty, Evolution, and Economic Theory’. *Journal of Political Economy* 58, no. 3 (June 1950): 211–21.

Antognazza, Maria Rosa. *The Oxford Handbook of Leibniz*. Oxford, United Kingdom: Oxford University Press, 2018.

Aristotle. *On Rhetoric: A Theory of Civic Discourse*. Translated by George A. Kennedy. Second Edition. Oxford, United Kingdom: Oxford University Press, 2007.

———. *Prior Analytics*. Translated by A.J. Jenkinson, 350 AD. http://classics.mit.edu/Aristotle/prior.html.

Arnauld, Antoine, and Pierre Nicole. *Logic or the Art of Thinking*. Translated by Jill Vance Buroker. 5th Edition. Cambridge Texts in the History of Philosophy. Cambridge University Press, 1996.

Avishai, Bernard. ‘The Pandemic Isn’t a Black Swan but a Portent of a More Fragile System’. *New Yorker*, 21 April 2020. https://www.newyorker.com/news/daily-comment/the-pandemic-isnt-a-black-swan-but-a-portent-of-a-more-fragile-global-system.

Azeredo, Francisco. ‘The Equity Premium: A Deeper Puzzle’. *Annals of Finance* 10 (19 April 2014): 347–73.

Bansal, Ravi, and Wilbur John Coleman II. ‘A Monetary Explanation of the Equity Premium, Term Premium, and Risk-Free Rate Puzzles’. *Journal of Political Economy* 104, no. 6 (December 1996): 1135–71. https://doi.org/10.1086%2F262056.

Barclay, Scott. ‘Handbook for Decisions Analysis’. McLean, Virginia: Decisions and Designs, Inc., 1977. https://files.eric.ed.gov/fulltext/ED153329.pdf.

Bayes, Thomas. ‘An Essay towards Solving a Problem in the Doctrine of Chances’. *Philosophical Transactions of the Royal Society of London* 53 (1763): 370–418. https://doi.org/10.1098/rstl.1763.0053.

Beinhocker, Eric D. ‘Reflexivity, Complexity, and the Nature of Social Science’. *Journal of Economic Methodology* 20, no. 4 (13 January 2014): 330–42. https://doi.org/10.1080/1350178X.2013.859403.

Benartzi, Shlomo, and Richard H. Thaler. ‘Myopic Loss Aversion and the Equity Premium Puzzle’. *The Quarterly Journal of Economics* 110, no. 1 (February 1995): 73–92. https://doi.org/10.2307/2118511.

Benello, Allen C., Michael van Biema, and Thomas Carlisle. *Concentrated Investing: Strategies of the World's Greatest Concentrated Value Investors*. Hoboken, New Jersey: John Wiley & Sons, Inc, 2016.

Berkeley, George. *A Treatise Concerning the Principles of Human Knowledge*. Project Gutenberg, 2009. https://www.gutenberg.org/files/4723/4723-h/4723-h.htm.

Bernoulli, Daniel. 'Exposition of a New Theory on the Measurement of Risk'. *Econometrica* 22, no. 1 (1 January 1954): 23–36.

Bernoulli, Jacob. *The Art of Conjecturing: Together with Letter to a Friend on Sets in Court Tennis*. Translated by Edith Dudley Sylla. Baltimore: Johns Hopkins University Press, 2005.

Bernstein, Peter L. *Against the Gods: The Remarkable Story of Risk*. New York: John Wiley & Sons, Inc, 1996.

Bessembinder, Hendrik. 'Do Stocks Outperform Treasury Bills?' *Journal of Financial Economics*, 28 May 2018. https://doi.org/10.2139/ssrn.2900447.

Bezemer, Dirk J. '"No One Saw This Coming": Understanding Financial Crisis Through Accounting Models'. *MPRA Paper No. 15892*, 16 June 2009. https://mpra.ub.uni-muenchen.de/15892/1/MPRA_paper_15892.pdf.

Bishop, Christopher M. *Pattern Recognition and Machine Learning*. First Edition. Springer, 2006.

Blikre, Jared. 'How Warren Buffett's Airline Stocks Have Performed since Berkshire Hathaway Sold Them'. Yahoo! Finance. *Yahoo! Finance* (blog), 27 April 2021. https://finance.yahoo.com/news/how-warren-buffetts-airline-stocks-have-performed-since-berkshire-hathaway-sold-them-134849843.html?guccounter=1&guce_referrer=aHR0cHM6Ly9kdWNrZHVja2dvLmNvbS8&guce_referrer_sig=AQAAALGsjkkuoTwfLNISYi7-_eQtYdUqPLjtvTKbiealr46tnvgfzKsT33hKIEzBjcgkHNmtvPjk1bKulgfZ6fCUH-FIwUNK1RZ3E77__wa1AYJiv0xAU-MFKdNdRdIPCntNYo5Q8sqiScIXw60dFjZ61WgNF3tq1ufroMO66idQcHwm.

Booth, David G., and Eugene F. Fama. 'Diversification Returns and Asset Contributions'. *Financial Analysts Journal* 48, no. 3 (June 1992): 26–32.

Born, Max. *Natural Philosophy of Cause and Chance: Being the Waynflete Lectures Delivered in the College of St. Mary Magdalen, Oxford, in Hilary Term 1948*. Oxford, United Kingdom: The Clarendon Press, 1949. https://archive.org/details/naturalphilosoph032159mbp.

Breiman, Leo. 'Investment Policies for Expanding Businesses Optimal in a Long-Run Sense'. *Stochastic Optimization Models in Finance*, 1975, 593–98. https://doi.org/10.1016/B978-0-12-780850-5.50050-2.

Bronk, Richard. 'Reflexivity Unpacked: Performativity, Uncertainty and Analytical Monocultures'. *Journal of Economic Methodology* 20, no. 4 (13 January 2014): 343–49. https://doi.org/10.1080/1350178X.2013.859404.

Caldwell, Bruce. 'George Soros: Hayekian?' *Journal of Economic Methodology* 20, no. 4 (13 January 2014): 350–56. https://doi.org/10.1080/1350178X.2013.859405.

Carnap, Rudolf. *Logical Foundations of Probability*. Second Edition. Chicago: The University of Chicago Press, 1962.

Gregory Chaitin, 'Irreducible Complexity in Pure Mathematics', *ArXiv* arXiv:math/0411091 (4 November 2004), https://doi.org/10.48550/arXiv.math/0411091.

Coase, Ronald H. 'The Nature of the Firm'. *Economica* 4, no. 16 (November 1937): 386–405. https://doi.org/10.2307/2626876.

Cohen, Edward E. *Athenian Economy and Society: A Banking Perspective*. Princeton, New Jersey: Princeton University Press, 1992.

Cox, Richard Threlkeld. 'Probability, Frequency and Reasonable Expectation'. *American Journal of Physics* 14, no. 1 (1946): 1–10. https://doi.org/10.1119/1.1990764.

———. *The Algebra of Probable Inference*. Baltimore, MD: Johns Hopkins University Press, 1961.

Copley, Bridget, and Fabienne Martin. *Causation in Grammatical Structures*. Oxford Studies in Theoretical Linguistics. Oxford, United Kingdom: Oxford University Press, 2015.

Cross, Rod, Harold Hutchinson, Harbir Lamba, and Doug Strachan. 'Reflections on Soros: Mach, Quine, Arthur and Far-from-Equilibrium Dynamics' 20, no. 4 (13 January 2014): 357–67. https://doi.org/10.1080/1350178X.2013.859406.

Crowe, Robert M., and Horn Ronald C. 'The Meaning of Risk'. *The Journal of Risk and Insurance* 34, no. 3 (September 1967): 459–74. https://doi.org/10.2307/250861.

Damodaran, Aswath. *Applied Corporate Finance*. Fourth Edition. John Wiley & Sons, Inc, 2014.

Davis, John B. 'Soros's Reflexivity Concept in a Complex World: Cauchy Distributions, Rational Expectations, and Rational Addiction'. *Journal of Economic Methodology* 20, no. 4 (13 January 2014): 368–76. https://doi.org/10.1080/1350178X.2013.859407.

De Ste Croix, Geoffrey E. M. 'Ancient Greek and Roman Maritime Loans'. In *Debits, Credits, Finance and Profits: Essays in Honour of W. T. Baxter*, edited by Harold C. Edey and Basil S. Yamey, 41–59. London, United Kingdom: Sweet and Maxwell, 1974.

Demosthenes. 'Against Lacritus'. Demosthenes 35, n.d. https://stuff.mit.edu/afs/athena/course/21/21h.401/www/local/demosthenes_35.html.

Devlin, Keith. *The Unfinished Game: Pascal, Fermat, and the Seventeenth-Century Letter That Made the World Modern*. Basic Books, 2010.

Dimson, Elroy, Paul Marsh, and Mike Staughton. 'The Worldwide Equity Premium: A Smaller Puzzle'. In *Handbook of the Equity Risk Premium*, edited by Rajnish Mehra, 609. Amsterdam, Netherlands: Elsevier, 2008.

Dixon, Huw. *Foundations of Economic Thought*. Edited by John Creedy. Oxford, United Kingdom: Basil Blackwell, 1990.

Donini, Francesco Maria, Maurizio Lenzerini, Daniele Nardi, Fiora Pirri, and Marco Schaerf. 'Nonmonotonic Reasoning'. *Artificial Intelligence Review* 4 (September 1990): 163–210.

Duhem, Pierre. 'Physical Theory and Experiment'. In *Philosophy of Science: The Central Issues*, edited by Martin Curd, J.A. Cover, and Christopher Pincock, Second., 257–79. New York: W.W. Norton & Company, 1998. https://personal.lse.ac.uk/robert49/teaching/methodologies/pdf/Duhem.pdf.

Dutka, Jacques. 'On the St. Petersburg Paradox'. *Archive for History of Exact Sciences* 39 (1988): 13–39.

Edwards, A.W.F. 'Pascal and the Problem of Points'. *International Statistical Review / Revue Internationale de Statistique*50, no. 3 (December 1982): 259–66. https://doi.org/10.2307/1402496.

———. 'Pascal's Problem: The "Gambler's Ruin"'. *International Statistical Review / Revue Internationale de Statistique*51, no. 1 (April 1983): 73–79. https://doi.org/10.2307/1402732.

Ellsberg, Daniel. 'Risk, Ambiguity, and the Savage Axioms'. *Quarterly Journal of Economics* 75, no. 4 (November 1961): 643–69. https://doi.org/10.2307/1884324.

Fama, Eugene F., and Kenneth R. French. 'Dissecting Anomalies'. *The Journal of Finance* 63, no. 4 (August 2008): 1653–78.

Fan, Steven, Scott Opsal, and Linda Yu. 'Equity Anomalies and Idiosyncratic Risk Around the World'. *Multinational Finance Journal* 19, no. 1 (28 May 2015): 33–75.

Feller, William. *An Introduction to Probability Theory and Its Applications*. Revised Edition. Vol. Vol. 1. Wiley Series in Probability and Mathematical Statistics. New York: John Wiley & Sons, Ltd, 1966.

Fienberg, Stephen E. 'When Did Bayesian Inference Become "Bayesian"?' *Bayesian Analysis* 1, no. 1 (2006): 1–40. https://doi.org/10.1214/06-BA101.

Finetti, Bruno de. 'La Prévision : Ses Lois Logiques, Ses Sources Subjectives'. *Annales de l'Institut Henri Poincaré* 7, no. 1 (1937): 1–68.

———. *Probability, Induction and Statistics;: The Art of Guessing*. Wiley Series in Probability and Mathematical Statistics. New York: John Wiley & Sons, 1972. 978-0471201403.

———. *Theory of Probability: A Critical Introductory Treatment*. Translated by Antonio Machí and Adrian Smith. Wiley Series in Probability and Statistics. John Wiley & Sons Ltd, 2017.

Fisher, Ronald. 'On the Mathematical Foundations of Theoretical Statistics.' *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 222 (1922): 309–68.

Franklin, James. 'Resurrecting Logical Probability'. *Erkenntnis* 55, no. 2 (1994): 277–305. https://doi.org/10.1023/A:1012918016159.

———. *The Science of Conjecture: Evidence and Conjecture Before Pascal*. Baltimore: Johns Hopkins University Press, 2015.

Frege, Gottlob. *The Frege Reader*. Edited by Michael Beaney. Blackwell Publishers, 1997.

Gabbay, Dov M., John Woods, and Stephan Hartmann. *Inductive Logic*. First Edition. Vol. Volume 10. 11 vols. Handbook of the History of Logic. Oxford, United Kingdom: North-Holland Publishing Company, 2011.

Gigerenzer, Gerd. 'The Bias Bias in Behavioral Economics'. *Review of Behavioral Economics* 5, no. 3–4 (31 December 2018): 303–36. https://doi.org/10.1561/105.00000092.

Gödel, Kurt. *On Formally Undecidable Propositions of Principia Mathematica and Related Systems I*. Translated by B. Meltzer. New York: Dover Publications, Inc., 1992. https://monoskop.org/images/9/93/Kurt_Gödel_On_Formally_Undecidable_Propositions_of_Principia_Mathematica_and_Related_Systems_1992.pdf.

Graham, Benjamin. *The Intelligent Investor: A Book of Practical Counsel*. Revised. Harper Business, 2006.

Graham, Benjamin, and David L. Dodd. *Security Analysis: Principles and Technique*. Sixth. New York: McGraw-Hill, 2009.

Graham, John R., and Campbell R. Harvey. 'The Equity Risk Premium in January 2007: Evidence from the Global Cfo Outlook Survey'. *SSRN*, 28 January 2007, 1–20. https://doi.org/10.2139/ssrn.959703.

Grant, Simon, and John Quiggin. 'The Risk Premium for Equity: Implications for Resource Allocation, Welfare and Policy'. *Risk & Uncertainty Program Working*

Paper: 8/R04, Risk and Sustainable Management Group Working Papers, 2004, 1–26. https://doi.org/10.22004/ag.econ.151167.

Green, Brett S., and Jeffrey Zwiebel. 'The Hot Hand Fallacy: Cognitive Mistakes or Equilibrium Adjustments? Evidence from Baseball'. *Stanford Graduate School of Business Working Papers* Working Paper No. 3101 (November 2013). https://www.gsb.stanford.edu/faculty-research/working-papers/hot-hand-fallacy-cognitive-mistakes-or-equilibrium-adjustments.

Hacking, Ian. 'Franklin's Conjecture'. *Isis* 95, no. 3 (September 2004): 460–64. https://doi.org/10.1086/428972.

———. *The Emergence of Probability: A Philosophical Study of Early Ideas about Probability, Induction and Statistical Inference*. Second Edition. Cambridge Series on Statistical & Probabilistic Mathematics. Cambridge, United Kingdom: Cambridge University Press, 2006.

Hannam, James. *God's Philosophers: How the Medieval World Laid the Foundations of Modern Science*. London: Icon Books, 2010.

Hayek, F.A. 'The Results of Human Action but Not of Human Design'. In *The Market and Other Orders*, edited by Bruce Caldwell, Volume 15:293–303. The University of Chicago Press, 2014.

Hoover, Calvin B. 'The Sea Loan in Genoa in the Twelfth Century'. *The Quarterly Journal of Economics* 40, no. 3 (May 1926): 495–529. https://doi.org/10.2307/1885175.

Hoppe, Hans-Hermann. *Economic Science and the Austrian Method*. Auburn, Alabama: Ludwig von Mises Institute, 2007. https://cdn.mises.org/Economic%20Science%20and%20the%20Austrian%20Method_3.pdf.

Hossenfelder, Sabine. *Existential Physics: A Scientist's Guide to Life's Biggest Questions*. Viking, 2022.

Howson, Colin. 'The Logic of Bayesian Probability'. In *Foundations of Bayesianism*, edited by David Corfield and Jon Williamson, 137–59. Applied Logic Series 24. Springer Netherlands, 2001.

Hume, David. *A Treatise of Human Nature: Being an Attempt to Introduce the Experimental Method of Reasoning into Moral Subjects*. Selby-Bigge, 1739. https://en.wikisource.org/wiki/Treatise_of_Human_Nature.

———. *An Enquiry Concerning Human Understanding*. London: A. Millar, 1748. https://books.google.fr/books?id=LB4VAAAAQAAJ&q=inauthor:%22David+Hume%22&pg=PR6&redir_esc=y#v=onepage&q=inauthor%3A%22David%20Hume%22&f=false.

Huygens, Christiaan. *De Ratiociniis in Ludo Aleae*. London, United Kingdom: S. Keimer, 1714. https://math.dartmouth.edu/~doyle/docs/huygens/huygens.pdf.

Jaynes, E.T. *Probability Theory: The Logic of Science*. Edited by G. Larry Bretthorst. Cambridge, United Kingdom: Cambridge University Press, 2003.

Jensen, Michael C. 'Maximization, Stakeholder Theory, and the Corporate Objective Function'. *Business Ethics Quarterly* 12, no. 2 (April 2002): 235–56.

Julliard, Christian, and Anisha Ghosh. 'Can Rare Events Explain the Equity Premium Puzzle?' *The Review of Financial Studies* 25, no. 10 (October 2012): 3037–76.

Juvenal. *D. IVNI IVVENALIS SATVRA VI*. Saturae, late 1st or early 2nd century. http://www.thelatinlibrary.com/juvenal/6.shtml.

———. *Satire VI*. Satires, late 1st or early 2nd century. https://sourcebooks.fordham.edu/ancient/juvenal-satvi.asp.

Kahneman, Daniel. 'Maps of Bounded Rationality'. *The American Economic Review* 93, no. 5 (December 2003): 1449–75.

Kahneman, Daniel, and Amos Tversky. 'Prospect Theory: An Analysis of Decision under Risk'. *Econometrica* 47, no. 2 (March 1979): 263–92. https://doi.org/10.2307/1914185.

Kelly, Jr., John Larry. 'A New Interpretation of Information Rate'. *The Bell System Technical Journal* 35, no. 4 (July 1956): 917–26. https://doi.org/10.1002/j.1538-7305.1956.tb03809.x.

Kent, Sherman. *Words of Estimative Probability*. Langley, Virginia: CIA Historical Review Program, 1964.

Keynes, John Maynard. *A Treatise on Probability*. The Collected Writings of John Maynard Keynes, Volume III. Cambridge, United Kingdom: Cambridge University Press for the Royal Economic Society, 2013.

———. 'F.P. Ramsey'. In *Essays in Biography*, edited by Austin Robinson and Donald Moggridge, Fifth Edition., Volume X:335–46. The Collected Writings of John Maynard Keynes. Cambridge, United Kingdom: Cambridge University Press for the Royal Economic Society, 2013.

———. 'The General Theory of Employment'. *The Quarterly Journal of Economics* 51, no. 2 (February 1937): 209–23. https://doi.org/10.2307/1882087.

———. *The General Theory of Employment, Interest, and Money*. First Edition. Cham, Switzerland: Palgrave Macmillan, 2018.

Knight, Frank H. *Risk, Uncertainty and Profit*. Reprints of Economic Classics. New York: Augustus M. Kelley, Bookseller, 1964. https://mises.org/library/risk-uncertainty-and-profit.

Kocherlakota, Narayana R. ‘The Equity Premium: It’s Still a Puzzle’. *Journal of Economic Literature* 34, no. 1 (March 1996): 42–71.

Koller, Tim, Marc Goedhart, and David Wessels. *Valuation: Measuring and Managing the Value of Companies*. Seventh Edition. Wiley Finance. Hoboken, New Jersey: John Wiley & Sons, Inc., 2020. https://www.wiley.com/go/valuation7e.

Koller, Tim, Dan Lovallo, and Zane Williams. ‘Overcoming a Bias against Risk’. McKinsey & Company. *Our Insights*(blog), 1 August 2012. https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/overcoming-a-bias-against-risk.

Lakatos, Imre. ‘Falsification and the Methodology of Scientific Research Programmes’. In *Criticism and the Growth of Knowledge: Proceedings of the International Colloquium in the Philosophy of Science*, edited by Imre Lakatos and Alan Musgrave. Cambridge, United Kingdom: Cambridge University Press, 1965.

———. *Proofs and Refutations: The Logic of Mathematical Discovery*. Edited by John Worrall and Elie Zahar. Cambridge Philosophy Classics. Cambridge, United Kingdom: Cambridge University Press, 2015.

———. *The Methodology of Scientific Research Programmes: Volume 1: Philosophical Papers*. Cambridge, United Kingdom: Cambridge University Press, 1980.

Laplace, Pierre-Simon. *Philosophical Essay on Probabilities*. Translated by Andrew Dale. Sources in the History of Mathematics and Physical Sciences 13. New York: Springer-Verlag New York, Inc., 1995.

Latané, Henry Allen. ‘Criteria for Choice Among Risky Ventures’. *Journal of Political Economy* 67, no. 2 (April 1959): 144–55. https://doi.org/10.1086/258157.

Lefebvre, Vladimir A. *Algebra of Conscience: Revised Edition with a Second Part with a New Foreword by Anatol Rapoport*. 1st ed. Theory and Decision Library 26. Netherlands: Springer Netherlands, 1982.

Leibniz, Gottfried Wilhelm. *New Essays Concerning Human Understanding*. Chicago: Open Court Publishers, 1916. https://archive.org/details/newessaysconcern01leib.

Levine, Steve. ‘Farm Bankruptcies Shed New Light on Perils of Big Agriculture’. Axios. *Axios* (blog), 31 March 2019. https://www.axios.com/iowa-country-of-monopolies-7de113b7-2860-4fa2-83ea-32c5b6766c15.html.

LIGO Scientific Collaboration, and Virgo Collaboration. ‘Observation of Gravitational Waves from a Binary Black Hole Merger’. *Physical Review Letters* 116, no. 061102 (12 February 2016): 1–16. https://doi.org/10.1103/PhysRevLett.116.061102.

Long, Roderick T. ‘Anti-Psychologism in Economics: Wittgenstein and Mises’. *The Review of Austrian Economics* 17, no. 4 (2004): 345–69.

Maanen, Jan van. ‘Teaching Geometry to 11 Year Old “Medieval Lawyers”’. *The Mathematical Gazette*, The Use of the History of Mathematics in the Teaching of Mathematics, 76, no. 475 (March 1992): 37–45. https://doi.org/10.2307/3620375.

Mackie, John L. *The Cement of the Universe: A Study of Causation*. Clarendon Library of Logic and Philosophy. Oxford, United Kingdom: Clarendon Paperbacks, 1980.

Malkiel, Burton G. *A Random Walk Down Wall Street*. Eleventh Edition. New York: W.W. Norton & Company, 2015.

Mao, James C.T. ‘Survey of Capital Budgeting: Theory and Practice’. *Journal of Finance* 25, no. 2 (May 1970): 349–60. https://doi.org/10.1111/j.1540-6261.1970.tb00513.x.

Markowitz, Harry. ‘Portfolio Selection’. *The Journal of Finance* 7, no. 1 (March 1952): 77–91. https://doi.org/10.2307/2975974.

Markowitz, Harry M. *Portfolio Selection: Efficient Diversification of Investments*. New York: John Wiley & Sons, Inc, 1959.

Marshall, Alfred. *Principles of Economics*. New York: Macmillan, 1890.

McIrvine, Edward C., and Myron Tribus. ‘Energy and Information’. *Scientific American* 224 (1 September 1971): 178–84.

Mehra, Rajnish. ‘The Equity Premium: Why Is It a Puzzle?’ *Financial Analysts Journal* 59, no. 1 (February 2003): 54–69.

Mehra, Rajnish, and Edward C. Prescott. ‘Non-Risk-Based Explanations of the Equity Premium’. In *Handbook of the Equity Risk Premium*, edited by Rajnish Mehra, 609. Amsterdam, Netherlands: Elsevier, 2008.

———. ‘The Equity Premium: A Puzzle’. *Journal of Monetary Economics* 15, no. 2 (March 1985): 145–61. https://doi.org/10.1016/0304-3932(85)90061-3.

———. ‘The Equity Premium in Retrospect’. *NBER Working Paper No. W9525*, 4 March 2003, 1–78. https://doi.org/10.3386/w9525.

Melo, Wolfgang David Cirilo de, and James Cussens. ‘Leibniz on Estimating the Uncertain: An English Translation of De Incerti Aestimatione with Commentary’. *The Leibniz Review* 14 (2004): 31–41.

Miller, John H., and Scott E. Page. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*. Illustrated Edition. Princeton Studies in Complexity. Princeton University Press, 2007.

Miller, Joshua B., and Adam Sanjurjo. ‘Is It a Fallacy to Believe in the Hot Hand in the NBA Three-Point Contest?’ *European Economic Review* 138 (September 2021). https://doi.org/10.1016/j.euroecorev.2021.103771.

———. ‘Surprised by the Gambler’s and Hot Hand Fallacies? A Truth in the Law of Small Numbers’. *Econometrica* 86, no. 6 (17 December 2018): 2019–47. https://doi.org/10.2139/ssrn.2627354.

Millett, Paul. ‘Maritime Loans and the Structure of Credit in Fourth-Century Athens’. In *Trade in the Ancient Economy*, edited by Peter Garnsey, Keith Hopkins, and C. R. Whittaker, 36–52. Berkeley and Los Angeles: University of California Press, 1983.

Mises, Ludwig von. *Economic Calculation in the Socialist Commonwealth*. Translated by S. Adler. Auburn, Alabama: Mises Institute, 2012. https://cdn.mises.org/economic_calculation_in_the_socialist_commonwealth.pdf.

———. *Human Action: A Treatise on Economics*. The Scholar’s Edition. Auburn, Alabama: Ludwig von Mises Institute, 1998.

Mises, Richard von. *Probability, Statistics and Truth*. Translated by Hilda Geiringer. Second Revised English Edition. New York: Dover Publications, Inc., 1981.

Moeller, Axel. *Alternative Initial Public Offering Models: The Law and Economics Pertaining to Shell Company Listings on German Capital Markets*. Studien Zum Ausländischen und Internationalen Privatrecht. Mohr Siebeck, 2017.

Montmort, Rémond de. *Essay d’analyse Sur Les Jeux de Hazard*. Seconde Édition. Jacque Quillau, 1713. https://gallica.bnf.fr/ark:/12148/bpt6k110519q.image.

Neumann, John von, and Oskar Morgenstern. *Theory of Games and Economic Behaviour*. Third Edition. Princeton: Princeton University Press, 1953.

Pascal, Blaise. *Les Provinciales*. Mozambook, 2001.

———. *Traité du triangle arithmétique avec quelques autres petits traitez sur la même matière*. Paris, France: Guillaume Desprez, 1665. https://gallica.bnf.fr/ark:/12148/btv1b86262012.texteImage.

Pathria, R.K., and Paul D. Beale. *Statistical Mechanics*. Fourth Edition. Academic Press, 2022.

Peters, Ole. ‘Comment on D. Bernoulli (1738)’. *Researchers.One*, 17 June 2020. https://researchers.one/articles/19.05.00001.

———. ‘The Ergodicity Problem in Economics’. *Nature Physics* 15 (2 December 2019): 1216–21. https://doi.org/10.1038/s41567-019-0732-.

———. ‘The Time Resolution of the St. Petersburg Paradox’. *Philosophical Transactions of the Royal Society A* 369, no. 1956 (13 December 2011): 4913--4931. https://doi.org/10.1098/rsta.2011.0065.

Peters, Ole, and Alexander Adamou. ‘Ergodicity Economics’, 30 June 2018.

Peters, Ole, and Murray Gell-Mann. ‘Evaluating Gambles Using Dynamics’. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 26, no. 023103 (2 February 2016): 1–11. https://doi.org/10.1063/1.4940236.

Pollock, John L. ‘How Do You Maximize Expectation Value?’ 17, no. 3 (September 1983): 409–21. https://doi.org/10.2307/2215257.

———. ‘Should We Maximize Expected Value?’ Journal article, 29 September 2010. https://web.archive.org/web/20100629151657/http://oscarhome.soc-sci.arizona.edu/ftp/PAPERS/MEV.ps.gz.

Pólya, George. *Mathematics and Plausible Reasoning*. Second Edition. Vol. Vol. II. Patterns of Plausible Inference. Princeton, New Jersey: Princeton University Press, 1968.

Popper, Karl. ‘Replies to My Critics’. In *The Philosophy of Karl Popper*, edited by Paul A. Schilpp, 961–1200. Library of Living Philosophers. La Salle, Illinois: Open Court Publishers, 1974.

———. *The Logic of Scientific Discovery*. Abingdon-on-Thames: Routledge, 1959.

Poundstone, William. *Fortune’s Formula: The Untold Story of the Scientific Betting System That Beat the Casinos and Wall Street*. New York: Hill and Wang, 2005.

Puhvel, Jaan. ‘The Origin of Etruscan Tusna (“Swan”)’. *The American Journal of Philology* 105, no. 2 (Summer 1984): 209–12. https://doi.org/10.2307/294875.

Quine, Willard Van Orman. ‘Two Dogmas of Empiricism’. *The Philosophical Review* 60, no. 1 (1951): 20–43.

Ramsey, Frank P., and Richard B. Braithwaite. ‘Truth and Probability’. In *The Foundations of Mathematics and Other Logical Essays*, 156–98. London, United Kingdom: Kegan, Paul, Trench, Trubner & Co., 1926. https://core.ac.uk/download/pdf/7048428.pdf.

Rietz, Thomas A. ‘The Equity Risk Premium a Solution’. *Journal of Monetary Economics* 22, no. 1 (July 1988): 117–31. https://doi.org/10.1016/0304-3932(88)90172-9.

Robinson, Joan. *Economics Is a Serious Subject: The Apologia of an Economist to the Mathematician, the Scientist and the Plain Man*. Cambridge, United Kingdom: W. Heffer & Sons Ltd., 1932.

Roubini, Nouriel, and Stephen Mihm. *Crisis Economics: A Crash Course in the Future of Finance*. New York: Penguin Books, 2011.

Roy, A.D. ‘Safety First and the Holding of Assets’. *Econometrica* 20, no. 3 (July 1952): 431–49. https://doi.org/10.2307/1907413.

Russell, Bertrand. ‘A Treatise on Probability. By John Maynard Keynes.’ *The Mathematical Gazette* 32, no. 300 (July 1948): 52–159. https://doi.org/10.2307/3609931.

Ryan, Leo V., F. Byron Nahser, and Wojciech Gasparski, eds. *Praxiology and Pragmatism: Praxiology: The International Annual of Practical Philosophy and Methodology*. Vol. 10. New Brunswick, New Jersey: Transaction Publishers, 2002. https://archive.org/details/praxiologypragma0010unse.

Sambursky, Shmuel. *The Physical World of the Greeks*. Translated by Merton Dagut. Princeton, New Jersey: Princeton University Press, 2014.

Samuelson, Paul A. 'Lifetime Portfolio Selection by Dynamic Stochastic Programming'. *The Review of Economics and Statistics* 51, no. 3 (August 1969): 239–46. https://doi.org/10.1016/B978-0-12-780850-5.50044-7.

———. 'The "Fallacy" of Maximizing the Geometric Mean in Long Sequences of Investing or Gambling'. *Proceedings of the National Academy of Sciences* 68, no. 10 (October 1971): 2493–96.

Shackle, G.L.S. *Expectations, Investment and Income*. Oxford, United Kingdom: Clarendon Press, 1968.

Shakespeare, William. *Hamlet: Prince of Denmark*. Edited by Philip Edwards. Updated Edition. The New Cambridge Shakespeare. Cambridge, United Kingdom: Cambridge University Press, 2003.

Shannon, Claude L. 'The Mathematical Theory of Communication'. In *The Mathematical Theory of Communication*, 16th Printing edition., 144. Urbana and Chicago: University of Illinois Press, 1971.

Shleifer, Andrei, and Robert W. Vishny. 'The Limits of Arbitrage'. *The Journal of Finance* 52, no. 1 (March 1997): 35–55. https://doi.org/10.2307/2329555.

Shuler, Matt. 'Buffett Is Shortsighted on Airlines: LUV Is A Steal at This Price'. New Constructs. *New Constructs* (blog), 4 May 2020. https://www.newconstructs.com/buffett-is-shortsighted-on-airlines-luv-is-a-steal-at-this-price/.

Soklakov, Andrei N. 'One Trade at a Time - Unravelling the Equity Premium Puzzle'. *SSRN*, 5 August 2015, 1–21. https://doi.org/10.2139/ssrn.2639388.

Sornette, Didier. 'Dragon-Kings, Black Swans and the Prediction of Crises'. *Swiss Finance Institute Research Paper*, Swiss Finance Institute Research Paper Series, no. 09–36 (8 September 2009): 1–21. https://doi.org/10.2139/ssrn.1470006.

Soros, George. 'Fallibility, Reflexivity, and the Human Uncertainty Principle'. *Journal of Economic Methodology* 20, no. 4 (2013): 309–29. https://doi.org/10.1080/1350178X.2013.859415.

———. *The Alchemy of Finance*. New York: Simon & Schuster, 1987.

Sortino, Frank A. *The Sortino Framework for Constructing Portfolios: Focusing on Desired Target Return to Optimize Upside Potential Relative to Downside Risk*. New York: Elsevier Finance, 2010.

Spanos, Aris. *Statistical Foundations of Econometric Modelling*. Cambridge, United Kingdom: Cambridge University Press, 1986.

Spitznagel, Mark. *Safe Haven Investing for Financial Storms*. Hoboken, New Jersey: John Wiley & Sons, Inc., 2021.

P.B. Stark and D.A. Freedman, 'What Is the Chance of an Earthquake?', in *Earthquake Science and Seismic Risk Reduction*, ed. Francesco Mulargia and Robert G. Geller, vol. 32, NATO Science Series, IV: Earth and Environmental Sciences (Dordrecht, The Netherlands: Kluwer Academic Publishers, 2016), 201–13, https://www.stat.berkeley.edu/~stark/Preprints/611.pdf.

Stigler, Stephen M. *The History of Statistics: The Measurement of Uncertainty Before 1900*. Illustrated Edition. Cambridge, Massachusetts, and London, England: The Belknap Press of Harvard University Press, 1986.

Strasser, Christian, and G. Aldo Antonelli. 'Non-Monotonic Logic'. In *Stanford Encyclopedia of Philosophy*, 20 April 2019. https://plato.stanford.edu/entries/logic-nonmonotonic/.

Taleb, Nassim Nicholas. *Skin in the Game: Hidden Asymmetries in Daily Life*. Incerto 5. New York: Random House, 2018.

———. *The Black Swan: The Impact of the Highly Improbable*. 2nd ed. London, United Kingdom: Penguin, 2010.

Taleb, Nassim Nicholas, and Dhruv Madeka. 'All Roads Lead to Quantitative Finance'. *Quantitative Finance* 19, no. 11 (August 2019): 1775–76. https://doi.org/10.1080/14697688.2019.1639803.

Tetlock, Philip E. *Expert Political Judgment: How Good Is It? How Can We Know?* New Edition. Princeton, New Jersey: Princeton University Press, 2017.

Tetlock, Philip E., and Dan Gardner. *Superforecasting: The Art and Science of Prediction*. New York: Crown Publishers, 2015.

Thomas, Timothy L. 'Russia's Reflexive Control Theory and the Military'. *Journal of Slavic Military Studies* 17 (2004): 237–56. https://doi.org/10.1080/13518040490450529.

Thorp, Edward O. 'Beat the Dealer: A Winning Strategy for the Game of Twenty-One', 220. Vintage, 1966.

———. 'Optimal Gambling Systems for Favorable Games'. *Review of the International Statistical Institute* 37, no. 3 (1969): 273–93.

———. 'Portfolio Choice and the Kelly Criterion'. *Stochastic Optimization Models in Finance*, 1971, 599–619. https://doi.org/10.1016/B978-0-12-780850-5.50051-4.

———. 'The Kelly Criterion in Blackjack, Sports Betting, and the Stock Market'. Montreal, 1997. https://doi.org/10.1016/S1872-0978(06)01009-X.

Van Horn, Kevin S. ‘Constructing a Logic of Plausible Inference: A Guide to Cox’s Theorem’. *International Journal of Approximate Reasoning* 34, no. 1 (September 2003): 3–24. https://doi.org/10.1016/S0888-613X(03)00051-3.

Venn, John. *The Logic of Chance: An Essay on the Foundations and Province of the Theory of Probability with Especial Reference to Its Logical Bearings and Its Application to Moral and Social Sciences and to Statistics*. 3rd Edition. London, United Kingdom: MacMillan and Co., 1888. https://www.gutenberg.org/files/57359/57359-pdf.pdf.

Wald, Abraham. ‘Contributions to the Theory of Statistical Estimation and Testing Hypotheses’. *Annals of Mathematical Statistics* 10, no. 4 (December 1939): 299–326. https://doi.org/10.1214/aoms/1177732144.

———. *Statistical Decision Functions*. Second Edition. Wiley Publications in Statistics. New York: John Wiley and Sons, Inc., 1950.

Weaver, Warren. ‘Some Recent Contributions to the Mathematical Theory of Communication’. In *The Mathematical Theory of Communication*, 16th Printing edition., 144. Urbana and Chicago: The University of Illinois Press, 1971.

Williams, John Burr. *The Theory of Investment Value*. Amsterdam, Netherlands: North-Holland Publishing Company, 1938.

Wucker, Michelle. *The Gray Rhino: How to Recognize and Act on the Obvious Dangers We Ignore*. New York: St. Martin’s Press, 2016.

Zuckerman, Gregory. *The Man Who Solved the Market: How Jim Simons Launched the Quant Revolution*. Portfolio / Penguin, 2019.

Table of Figures

Figure 1 : The Meaning of Words of Estimative Probability. Source: (Barclay 1977, 68) 58
Figure 2 : Expected Value and the Central Limit Theorem 70
Figure 3 : The Mean Payout Approaches Its Limit 71
Figure 4 : Huygens' Leads to Gambler's Ruin 73
Figure 5: Gambler's Ruin Across Many Lives 73
Figure 6 : Situation-value as a function of accumulated goods. Source: Pollock (2010) 76
Figure 7 : Life-accumulations. Source: Pollock (2010) 77
Figure 8: Capped Saint Petersburg Wager 79
Figure 9: The Logarithm of Wealth. Source: Bernoulli (1954) 83
Figure 10: Concavity of the Curve 88
Figure 11: Berkshire Hathaway's Airline Stocks Performance. Source: Bilkre (2021) 89
Figure 12: Entropy With Two Possible Outcomes 98

Table of Tables

Table 1: Converting Bernoulli's Notation 83