Catastrophic Risk, Rare Events, and Black Swans: Could There Be a Countably Additive Synthesis?

Peter Hammond

No 1060

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

Catastrophic Risk, Rare Events, and Black Swans: Could There Be a Countably Additive Synthesis?

Peter J. Hammond: p.j.hammond@warwick.ac.uk
Department of Economics
and CAGE (Centre for Competitive Advantage in the Global Economy)
University of Warwick, Coventry CV4 7AL, UK.

2015 May 1st, typeset from pjhForAFOSR2.tex

Abstract

Catastrophic risk, rare events, and black swans are phenomena that require special attention in normative decision theory. Several papers by Chichilnisky integrate them into a single framework with finitely additive subjective probabilities. Some precursors include: (i) following Jones-Lee (1974), undefined willingness to pay to avoid catastrophic risk; (ii) following Rényi (1955, 1956) and many successors, rare events whose probability is infinitesimal. Also, when rationality is bounded, enlivened decision trees can represent a dynamic process involving successively unforeseen "true black swan" events. One conjectures that a different integrated framework could be developed to include these three phenomena while preserving countably additive probabilities.

Acknowledgements

Generous support during 2007–10 from a Marie Curie Chair funded by the European Commission under contract number MEXC-CT-2006-041121 is gratefully acknowledged. So is the opportunity to present this work at the AFOSR (Air Force Office of Scientific Research) workshop on Catastrophic Risk, organized by Graciela Chichilnisky, and held at SRI, Menlo Park, California, on May 31st and June 1st, 2012. Some of the material on black swans is adapted from the earlier informal online publication Hammond (2009).

1 Introduction

1.1 Countably Additive Subjective Probability

Savage's (1954) classic, *The Foundations of Statistics*, provided an axiom system sufficient to imply that a decision maker's preferences could be represented by the expected value of a von Neumann–Morgenstern utility function, with personal or subjective probabilities attached to unknown events ranging over a sample space of states of the world. His axioms, however, implied that probabilities are only finitely rather than countably additive. Yet countable additivity is a key measure-theoretic property that probabilitists since the time of Kolmogorov (1933), at least, are accustomed to using. It allows, for instance, the probability of an interval of the real line to be found by integrating a density function over that interval.

1.2 Monotonicity

To bridge this gap between finite and countable additivity, Villegas (1964), Arrow (1965) and Fishburn (1982) all introduced an additional monotonicity axiom ensuring that subjective probabilities are countably additive. One version of the relevant axiom can be derived by combining slightly modified versions of two axioms set out in Sections 7.2 and 7.3 of Hammond (1998b).

Let $(Y, \mathcal{F})$ denote a measurable space of *consequences*, and $(S, \mathcal{S})$ a measurable space of *states of the world*. Following the evocative terminology introduced by Anscombe and Aumann (1963), let $\Delta(Y, \mathcal{F})$ denote the space of *roulette lotteries* in the form of probability measures over $(Y, \mathcal{F})$.

In the special case when S is a finite set, for each $E \subseteq S$, let Y^E denote the Cartesian product set $\prod_{s \in E} Y_s$, where each Y_s is a copy of Y, and let $\mathcal{F}^E$ denote the product σ-field $\bigotimes_{s \in E} \mathcal{F}_s$, where each $\mathcal{F}_s$ is a copy of $\mathcal{F}$. Consider then the space $\Delta(Y^E, \mathcal{F}^E)$ of roulette lotteries, in the form of probability measures over $(Y^E, \mathcal{F}^E)$ whose random outcomes are horse lotteries y^E in the space of measurable mappings from E to $(Y, \mathcal{F})$.

The key *reversal of order axiom* (RO) due to Anscombe and Aumann (1963) treats, for any event $E \subseteq S$, any pair $\lambda^E, \mu^E \in \Delta(Y^E, \mathcal{F}^E)$ as equivalent if and only if their marginal measures $\lambda_s, \mu_s \in \Delta(Y, \mathcal{F})$ are equal for each state $s \in E$. Then each $\pi^E \in \Delta(Y^E, \mathcal{F}^E)$ can be identified with the list $\langle \pi_s \rangle_{s \in E}$ of marginal probability measures $\pi_s \in \Delta(Y, \mathcal{F})$. In particular, this treats as irrelevant the extent of any correlation between consequences $y_s \in Y$ that arise in different states $s \in E$.

Next, we revert to the case of a general measurable space $(S, \mathcal{S})$. Then, for each measurable event $E \in \mathcal{S}$, define the conditional sub-σ-field

$$\mathcal{S}_{|E} := \{G \in \mathcal{S} \mid G \subseteq E\} \subseteq \mathcal{S}$$

Obviously, in case $E = S$, this definition implies that $\mathcal{S}_{|S} = \mathcal{S}$.

In the spirit of the case when S is a finite set, for each measurable event $E \in \mathcal{S}$, let $\Delta(Y^E, \mathcal{S}_{|E}, \mathcal{F})$ denote the space of functions $\pi^E : E \to \Delta(Y, \mathcal{F})$ with the property that, for each $K \in \mathcal{F}$, the mapping

$$E \ni s \mapsto \pi^E(s, K) \in \mathbb{R}_+$$

is measurable w.r.t. the σ-field $\mathcal{S}_{|E}$ on E and the Borel σ-field on $\mathbb{R}$.

The other axioms to be discussed here concern:

1. the preference ordering $\succsim^*$ on $\Delta(Y, \mathcal{F})$ having the property that for each $y \in Y$, in addition to the set $\{y\}$, the upper and lower contour sets
$$\{y' \in Y \mid \delta_{y'} \succsim^* \delta_y\} \quad \text{and} \quad \{y' \in Y \mid \delta_{y'} \precsim^* \delta_y\}$$
are both $\mathcal{F}$-measurable;

2. for each measurable event $E \in \mathcal{S}$, the conditional preference ordering $\succsim^E$ on $\Delta(Y^E, \mathcal{S}_{|E}, \mathcal{F})$.

Definition 1. ***Event Dominance (ED)***

Suppose that the event $E \in \mathcal{S}$, the list of probability measures $\pi^E = \langle \pi_s \rangle_{s \in E} \in \Delta(Y^E, \mathcal{S}_{|E}, \mathcal{F})$, and the simple lottery $\lambda \in \Delta(Y)$ are all given. Let $\lambda\, 1^E$ denote the particular list $\lambda^E = \langle \lambda_s \rangle_{s \in E} \in \Delta(Y^E, \mathcal{S}_{|E}, \mathcal{F})$ that satisfies $\lambda_s = \lambda$ for all $s \in E$. Then:

1. *$\pi_s \succsim^* \lambda$ (all $s \in E$) implies $\pi^E \succsim^E \lambda\, 1^E$;*

2. *$\pi_s \precsim^* \lambda$ (all $s \in E$) implies $\pi^E \precsim^E \lambda\, 1^E$.*

In case the set E is finite, condition (ED) is an obvious implication of Anscombe and Aumann's extension of Savage's sure thing principle. The force of (ED) comes in partially extending this principle to the case when E is any measurable subset of S.

Next, given any measurable event $E \in \mathcal{S}$ satisfying $\emptyset \neq E \neq S$, let $(\pi\, 1^E, \tilde{\pi}\, 1^{S \setminus E})$ denote the particular list of probability measures $\lambda^S = \langle \lambda_s \rangle_{s \in S} \in \Delta(Y^E, \mathcal{S}, \mathcal{F})$ whose marginal distribution $\lambda_s \in \Delta(Y, \mathcal{F})$ for each $s \in S$ is a roulette lottery that satisfies

$$\lambda_s = \begin{cases} \pi & \text{if } s \in E \\ \tilde{\pi} & \text{if } s \in S \setminus E \end{cases}$$

Definition 2. ***Event Continuity (EC)***
Let $\succsim^*$ on $\Delta(Y, \mathcal{F})$ and $\succsim^S$ on $\Delta(Y^S, \mathcal{S}^S, \mathcal{F})$ be fixed preference orderings. Suppose that the two measurable events $E, E^* \subset S$, as well as the sequence of measurable events E_k ($k \in \mathbb{N}$), and the two probability measures $\pi, \tilde{\pi} \in \Delta(Y, \mathcal{F})$, together satisfy:

1. $E_1 \subset E_2 \subset \ldots \subset E_k \subset E_{k+1} \subset \ldots \subset S$;
2. $E^* = \cup_{k=1}^{\infty} E_k$;
3. $\pi \succ^* \tilde{\pi}$;
4. $(\pi\, 1^{E^*}, \tilde{\pi}\, 1^{S \setminus E^*}) \succ^S (\pi\, 1^E, \tilde{\pi}\, 1^{S \setminus E})$.

Then there must exist a finite k such that $(\pi\, 1^{E_k}, \tilde{\pi}\, 1^{S \setminus E_k}) \succ^S (\pi\, 1^E, \tilde{\pi}\, 1^{S \setminus E})$.

Equivalently,

$$(\pi\, 1^{E_k}, \tilde{\pi}\, 1^{S \setminus E_k}) \succsim^S (\pi\, 1^E, \tilde{\pi}\, 1^{S \setminus E}) \text{ (all } k \in \mathbb{N}) \\ \Longrightarrow (\pi\, 1^{E^*}, \tilde{\pi}\, 1^{S \setminus E^*}) \succsim^S (\pi\, 1^E, \tilde{\pi}\, 1^{S \setminus E})$$

1.3 Beyond Monotonicity

In several recent papers, Chichilnisky (1996, 2000, 2009, 2010) has explored a particular weakening of this kind of monotonicity axiom. This weakening allows a revised decision theory in which rare events, catastrophes, perhaps even "black swans", can all be given more prominence. Of course, the weakening comes at the cost of allowing probabilities that are only finitely additive. For this reason, ultimately it may be useful to investigate whether some alternative approach could allow for such phenomena while retaining probabilities that are countably additive measures.

1.4 Outline of Paper

The rest of this paper considers three different strands of literature. First, Section 2 considers some background on the use of the word "catastrophe", in drama, mathematics, and finally decision theory. It goes on to formalize a notion of catastrophic risk in decision theory, based on pioneering work on the value of life due to Drèze (1962), followed by Jones-Lee (1974).

The second strand discussed in Section 3 concerns the use of infinitesimals to represent the subjective probability of events so rare that they should not be accorded any positive probability. Third, Section 4 offers a possible approach to modelling the "true black swans" that Taleb (2007) in particular regards as beyond any kind of systematic analysis. Finally, Section 5 combines a suggestion for an alternative synthesis of these three strands with some concluding remarks.

2 Catastrophic Risk

2.1 Etymology

According to `http://www.etymonline.com/`, the word "catastrophe" entered the English language during the 1530s with the meaning " 'reversal of what is expected' (especially a fatal turning point in a drama)". It is derived from the Greek "katastrophe", meaning "overturning; a sudden end", itself a compound of the prefix "kata" meaning "down" and "strephein" meaning "turn".

The extension of the meaning of "catastrophe" to include "sudden disaster" is first recorded in 1748. In medicine, catastrophe is often taken to mean death related to what should have been routine surgery. In engineering, a "catastrophic failure" is the complete breakdown of a system from which recovery is impossible. A celebrated example is the Tay Bridge disaster of 1879 which, thanks to William McGonagall's (1880) doggerel, has become a classic of British folklore.

There is a branch of mathematics known as "catastrophe theory" that concerns the possible instability of the minimum of a non-linear potential function when that function depends on exogenous parameters which may be subject to sudden shocks. The monograph by Thom (1972) provided a systematic classification of different types of catastrophe. Zeeman (1976) did much to popularize the application of catastrophe theory to the study of many different dynamic phenomena where there is a sudden change. These applications include:

- in animal psychology, aggression in dogs;
- in medicine, the beating heart;
- in structural engineering, beams that first buckle and then collapse;
- in economics and finance, crashes in stock markets, as well as Balasko's (1978) description of structural properties of the Walrasian equilibrium manifold.

2.2 Catastrophic Consequences

Standard decision theory considers acts whose consequences range over a specified *consequence domain* in the form of an abstract set Y equipped with a σ-algebra $\mathcal{F}$ of measurable sets. In principle, catastrophes can be described by letting the consequence domain Y be the union of the two disjoint measurable sets: (i) Y_0 of *non-catastropic* consequences; and (ii) Y_1 of *catastrophic* consequences.

Here, however, our concern will be to discuss how catastrophes can be modeled as events so extreme that a suitable money metric utility function becomes undefined whenever the probability of a catastrophe is sufficiently high. Accordingly, consider a *consequence domain* $K \times \mathbb{R}_+$ of pairs (κ, y) where:

1. $y \in \mathbb{R}_+$ is income or wealth (depending on context);
2. $\kappa \in K = \{0, 1\}$ is a binary indicator variable indicating whether a "catastrophe":
    - occurs, iff $\kappa = 1$;
    - or does not occur, iff $\kappa = 0$.

Hence $Y_0 = \{0\} \times \mathbb{R}_+$, whereas $Y_1 = \{1\} \times \mathbb{R}_+$,

Following Drèze (1962), consider too a consumer whose preference ordering $\succsim$ on the set $\Delta(K \times \mathbb{R}_+)$ of lotteries over $K \times \mathbb{R}_+$ is represented by the expected value $\mathbb{E}u$ of each real-valued von Neumann–Morgenstern utility function (or NMUF) $K \times \mathbb{R}_+ \mapsto (\kappa, y) \mapsto u(\kappa, y) \in \mathbb{R}$ in a unique cardinal equivalence class. The literature on decision theory inspired by Drèze often regards the mapping $y \mapsto u(\kappa, y)$ as a *state-dependent* utility function of income y, though it can perhaps be more usefully regarded as a *state-independent* utility function of the *fully specified consequence* (κ, y).

2.3 Assumptions

Within the framework of Section 2.2, we assume that:

1. for each fixed $\kappa \in K$, each NMUF $y \mapsto u(\kappa, y)$ is continuous, strictly increasing, and bounded above, with upper bound $\bar{u}_\kappa := \sup u(\kappa, y)$;

2. for each fixed $y \in \mathbb{R}_+$, one has $u(0, y) > u(1, y)$;

3. $\bar{u}_0 > \bar{u}_1$.

The second assumption, of course, is that the consumer is worse off with a catastrophe than without, *ceteris paribus*. Taking the limit as $y \to \infty$ implies that $\bar{u}_0 \geq \bar{u}_1$, obviously, but the third assumption that $\bar{u}_0 > \bar{u}_1$ strengthens this to a strict inequality. In particular, this third assumption holds if and only if there is a continuous extended utility function $\tilde{u} : K \times (\mathbb{R}_+ \cup \{\infty\}) \to \mathbb{R}$ for which there exists $y^* \in \mathbb{R}$ such that $\tilde{u}(0, y^*) = \tilde{u}(1, \infty)$ and so $\tilde{u}(0, y) > \tilde{u}(1, \infty)$ whenever $y > y^*$.

2.4 Money Metric Utility

Following Jones-Lee (1974), consider this consumer's willingness to pay for a reduction in the probability p of catastrophe. Specifically, consider any *reference* or *baseline* lottery

$$\lambda^R := (1 - p^R)\delta_{(0, y_0^R)} + p^R \delta_{(1, y_1^R)} \tag{1}$$

which is a mixture of the two degenerate lotteries $\delta_{(0, y_0^R)}$ and $\delta_{(1, y_1^R)}$ that attach probability one to the consequences $(0, y_0^R)$ and $(1, y_1^R)$ respectively. Thus, the consumer faces the probability p^R of a catastrophe, along with reference income levels y_κ^R ($\kappa \in \{1, 0\}$) with and without a catastrophe. Let

$$U^R := (1 - p^R) u(0, y_0^R) + p^R u(1, y_1^R) \tag{2}$$

denote expected utility in the reference situation. One can use these reference levels and the equation

$$(1-p)u(0,m) + pu(1,y_1) = U^R \tag{3}$$

in an attempt to define implicitly a *money metric* utility function

$$\mathbb{R}_+ \times [0,1] \ni (y_1; p) \mapsto m(y_1; p) \in \mathbb{R}_+ \tag{4}$$

Note that this function will be the same whenever u is replaced by an alternative NMUF that is cardinally equivalent.

Definition (4), when valid, implies that $m(y_1; p) - y_0^R$ is the consumer's *willingness to accept* the net increase $p - p^R$ in the risk of catastrophe, when compensation in the event of the catastrophe raises income from y_1^R to y_1. Alternatively, $y_0^R - m(y_1; p)$ is the consumer's (net) *willingness to pay*, in terms of foregone income in the absence of catastrophe, for the decrease in the probability of catastrophe from p^R to p.

2.5 A Critical Probability Level: Catastrophic Risk

The money metric utility function (4) really is defined by equation (3) for the pair $(y_1; p)$ if and only if

$$(1-p)u(0,0) + pu(1,y_1) \leq U^R.$$

Otherwise giving up all income is insufficient to compensate for the increase in p, which one could then regard as a *true* catastrophe.

In particular, the function (4) is defined iff $p \leq p_C$ for the *critical probability level* defined by

$$p_C := \frac{U^R - u(0,0)}{u(1,y_1) - u(0,0)} = \frac{(1-p^R)u(0,y_0^R) + p^R u(1,y_1^R) - u(0,0)}{u(1,y_1) - u(0,0)} \tag{5}$$

Thus, once p has reached p_C, no compensation is possible for any further increase in the probability of catastrophe.

Note that p_C, as the ratio of expected utility differences, is not only preserved under positive affine utility transformations. In addition, as discussed in Hammond (1998a), the formula (5) that expresses p_C as the ratio of utility differences implies that it must equal the constant marginal rate of substitution between shifts in probability away from $(0,0)$, the worst possible outcome without a catastrophe, toward respectively:

1. the reference lottery defined by (1);
2. the consequence $(1, y_1)$ that represents the occurrence of the catastrophe combined with the income level y_1.

2.6 Extreme Economic Catastrophes

One can also have an *extreme catastrophe* where p is large enough to satisfy

$$(1-p)u(0,0) + p\bar{u}_1 > U^R$$

This, of course, is equivalent to

$$p > \frac{U^R - u(0,0)}{\bar{u}_1 - u(0,0)} \tag{6}$$

Inequality (6) implies that the probability of catastrophe is so high that no matter how large y_1 may be, there is no value of m that satisfies (3). In this sense, compensation is completely impossible.

3 Rare Events

3.1 Standard Decision Theory

Standard decision theory uses the expected utility (EU) criterion. Traditionally, moreover, a distinction is made between *objective* and *subjective* EU theory, depending on whether one faces:

- *risk* or *roulette lotteries* described by *objective* probabilities, as in von Neumann and Morgenstern (1944) and then Jensen (1967);
- *uncertainty* or *horse lotteries* described by *subjective* probabilities, as in Savage (1954);
- combinations of roulette and horse lotteries, as in Anscombe and Aumann (1963).

3.2 Infinitesimal Probability

Recall that, by definition, an *infinitesimal* ϵ is some positive entity (not a real number) that is smaller than any positive real number in the sense that $0 < n\epsilon < 1$ for all natural numbers $n \in \mathbb{N}$. To accommodate rare events, one can follow the game-theoretic literature emanating from Selten (1975) by allowing "trembles" whose probability is taken to be some positive multiple of a particular *basic infinitesimal* ϵ. See Halpern (2009, 2010) for discussion of some recent developments.

3.3 Rare Events and Infinitesimal Probabilities

Probabilities must be:

1. *added* when calculating the probability of the union of two or more pairwise disjoint events;

2. *subtracted* when calculating the probability of the set-theoretic difference of any two events;

3. *multiplied* when compounding probabilities at successive stages of a stochastic process;

4. *divided* when calculating conditional probabilities.

This suggests that Selten's space of trembles should be enriched so that the extended probabilities we construct take values in an algebraic field, where all these four operations are well-defined — except, of course, when trying to divide by zero. This motivates the following definition:

Definition 3. *A* polynomial function *of* ϵ *takes the form*

$$P(\epsilon) \equiv \sum\nolimits_{k \in K} p_k \epsilon^k = \sum\nolimits_{j=1}^{r} p_{k_j} \epsilon^{k_j} \tag{7}$$

for some finite set $K = \{k_1, k_2, \ldots, k_r\} \subset \mathbb{Z}_+$*, where* $k_j < k_{j+1}$ *for* $j = 1, 2, \ldots, r-1$*, and* $p_k \neq 0$ *for all* $k \in K$*. The* leading non-zero coefficient *of the polynomial (7) is* p_{k_1}*. The polynomial (7) is* positive *just in case* $p_{k_1} > 0$.

A rational function *of* ϵ *takes the form of a quotient* $P(\epsilon)/Q(\epsilon)$ *of two polynomial functions of* ϵ*, where the denominator* $Q(\epsilon)$ *is positive. Without loss of generality, the leading non-zero coefficient of* $Q(\epsilon)$ *can be normalized to 1.*

Following Robinson (1973), define $\mathbb{R}(\epsilon)$ *as the algebraic field whose members are rational functions of* ϵ*, equipped with the standard algebraic binary operations of addition and multiplication, as well as the additive identity 0 and the multiplicative identity 1. Define the* positive cone $\mathbb{R}_+(\epsilon)$ *of rational functions* $P(\epsilon)/Q(\epsilon)$ *as those where* $P(\epsilon)$ *as well as* $Q(\epsilon)$ *is a positive polynomial.*

Following ideas that were surveyed in Hammond (1994), rare events E in a finite set S of states of the world can be modelled formally as having infinitesimal probability $p(E;\epsilon)$ in an extended EU theory with "non-Archimedean" probabilities in the positive cone $\mathbb{R}_+(\epsilon)$ of the field $\mathbb{R}(\epsilon)$. That is, we must have $p(E;\epsilon) = P(\epsilon)/Q(\epsilon)$ where the coefficient of ϵ^0 in the polynomial (7) is zero. Obviously one requires the probability mapping $2^S \ni E \mapsto p(E;\epsilon) \in \mathbb{R}_+(\epsilon)$ to satisfy the *additivity condition* $p(E;\epsilon) \equiv p(E';\epsilon) + p(E'';\epsilon)$ whenever $E = E' \cup E''$ with $E' \cap E'' = \emptyset$, as well as the *normalization condition* $p(S;\epsilon) \equiv 1$.

3.4 A Metric Completion

As discussed in Hammond (1999b), following an approach set out in Lightstone and Robinson (1975), the set $\mathbb{R}(\epsilon)$ of rational functions can be given a (real-valued) metric $d : \mathbb{R}(\epsilon) \times \mathbb{R}(\epsilon) \to \mathbb{R}_+$. This metric induces a very fine topology, according to which a sequence $r^{\mathbb{N}} = \langle r_n \rangle_{n\in\mathbb{N}}$ of real numbers converges to $r^* \in \mathbb{R}$ if and only if r_n is *eventually equal* to r^* — i.e., there exists $n^* \in \mathbb{N}$ such that $n \geq n^* \Longrightarrow r_n = r^*$.

Let $\mathbb{R}^{\mathbb{N}}(\epsilon)$ denote the Cartesian product of countably many copies of the algebraic field $\mathbb{R}(\epsilon)$. The elements of $\mathbb{R}^{\mathbb{N}}(\epsilon)$ are infinite sequences $r^{\mathbb{N}}(\epsilon) = (r^n(\epsilon))_{n\in\mathbb{N}}$ of rational functions of ϵ. Following standard terminology in metric space theory, say that $r^{\mathbb{N}}(\epsilon) = (r^n(\epsilon))_{n\in\mathbb{N}}$ is a *Cauchy sequence* if for every small $\delta > 0$, there exists $n_\delta \in \mathbb{N}$ such that whenever $n', n'' \in \mathbb{N}$ with $n' > n_\delta$ and $n'' > n_\delta$, one has $d(r^{n'}(\epsilon), r^{n''}(\epsilon)) < \delta$.

Define the binary relation $\sim$ on the space of Cauchy sequences in $\mathbb{R}^{\mathbb{N}}(\epsilon)$ so that $r^{\mathbb{N}}(\epsilon) \sim \tilde{r}^{\mathbb{N}}(\epsilon)$ just in case, for every small $\delta > 0$, there exists $n_\delta \in \mathbb{N}$ such that whenever $n', n'' \in \mathbb{N}$ with $n' > n_\delta$ and $n'' > n_\delta$, one has $d(r^{n'}(\epsilon), \tilde{r}^{n''}(\epsilon)) < \delta$. It is easy to check that the relation $\sim$ is symmetric, reflexive, and transitive — i.e., it is an *equivalence relation*. Then the metric space $(\mathbb{R}(\epsilon), d)$, like any other, has a *metric completion* consisting of equivalence classes of Cauchy sequences. In Hammond (1997) it is shown that each member of this metric completion can be expressed uniquely as a *power series* $\sum_{k=0}^{\infty} a_k \epsilon^k$ of the basic infinitesimal ϵ, for an infinite sequence $a^{\mathbb{N}} = (a_k)_{k \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ of real constants. We denote this metric completion by $(\mathbb{R}^\infty(\epsilon), d)$, where d denotes an obvious extension to the set $\mathbb{R}^\infty(\epsilon)$ of power series of the original metric d on the set $\mathbb{R}(\epsilon)$ of rational functions.

In the following, let $\mathbb{R}_+^\infty(\epsilon)$ denote the subset of power series that are *positive* in the sense that the leading non-zero coefficient is positive. We also introduce the *lexicographic strict ordering* $>_L$ on $\mathbb{R}^\infty(\epsilon)$, defined so that

$$\sum_{k=0}^{\infty} a_k \epsilon^k >_L \sum_{k=0}^{\infty} b_k \epsilon^k$$

if and only if the leading non-zero coefficient of the difference

$$\sum_{k=0}^{\infty} (a_k - b_k) \epsilon^k$$

is positive. Let $\geq_L$ denote the corresponding weak ordering defined so that

$$\sum_{k=0}^{\infty} a_k \epsilon^k \geq_L \sum_{k=0}^{\infty} b_k \epsilon^k \Longleftrightarrow \sum_{k=0}^{\infty} b_k \epsilon^k \not>_L \sum_{k=0}^{\infty} a_k \epsilon^k$$

3.5 Extended Probability Measures

In order to treat compound lotteries in decision trees where branches at one or more successive chance nodes can have infinitesimal probabilities, and also to have a satisfactory theory of subjective probability, it seems desirable to allow probabilities to have values in $\mathbb{R}_+^\infty(\epsilon)$ rather than just in $\mathbb{R}_+$.

Definition 4. *Let $(S, \mathcal{S})$ be any measurable state space S with σ-field $\mathcal{S}$. An* extended probability measure *on $(S, \mathcal{S})$ is a mapping*

$$\mathcal{S} \ni E \mapsto \pi(E; \epsilon) = \sum_{k=0}^{\infty} \pi_k(E) \epsilon^k \in \mathbb{R}^\infty(\epsilon)$$

that satisfies:

1. *$\pi(E; \epsilon) \in \mathbb{R}_+^\infty(\epsilon)$ for all $E \in \mathcal{S} \setminus \{\emptyset\}$;*
2. *$\pi(S; \epsilon) = 1$;*
3. *if the countable collection of sets E_n ($n \in \mathbb{N}$) is pairwise disjoint, then $\pi(\cup_n E_n; \epsilon) = \sum_n \pi(E_n; \epsilon)$ (countable additivity).*

Let $\Delta(S, \mathcal{S}; \mathbb{R}^\infty_+(\epsilon))$ denote the family of all extended probability measures on $(S, \mathcal{S})$.

Note that, apart from having values in the algebraic field $\mathbb{R}^\infty(\epsilon)$, such probabilities are required to be positive for all possible events; a zero probability is attached only to the empty set.

3.6 Extended Subjective Expected Utility

For the case when S is finite, Hammond (1997) offers axioms which imply that a preference ordering $\succsim$ over the space $\Delta(Y^S)$ of all possible combination of roulette and horse lotteries can be represented by the lexicographic weak ordering $\geq_L$ applied to subjectively expected utility, in the form of a power series

$$\sum\nolimits_{y^S \in Y^S} \lambda(y^S) \sum\nolimits_{s \in S} \pi(s; \epsilon) v(y_s) \in \mathbb{R}^\infty(\epsilon)$$

Note in particular that the von Neumann–Morgenstern utility function (or NMUF) $v : Y \to \mathbb{R}$ is *real* valued; there is no need for any form of lexicographic utility, as opposed to lexicographic expected utility. The following is the main theorem of Hammond (1997):

Theorem 1. *Let S denote a finite set of unknown states of the world, and Y a consequence domain. Suppose that all the seven axioms (O), (I*), (C*), (RO), (SI), (RC) and (XC) of Hammond (1997) are satisfied throughout the domain $\Delta(Y^S; \mathbb{R}^\infty_+(\epsilon))$ of consequence lotteries with non-Archimedean objective probabilities ranging over $\mathbb{R}^\infty_+(\epsilon)$. Unless there is universal indifference over the whole domain, there exist*

- *a unique extended subjective probability measure $p(\cdot; \epsilon)$ in the space $\Delta(S, \mathcal{S}; \mathbb{R}^\infty_+(\epsilon))$ of mappings $\mathcal{S} \ni E \mapsto p(E; \epsilon) \in \mathbb{R}^\infty_+(\epsilon)$;*
- *a unique cardinal equivalence class of real-valued NMUFs $v : Y \to \mathbb{R}$*

such that the preference ordering $\succsim^S$ on $\Delta(Y^S; \mathbb{R}^\infty(\epsilon))$ is represented by the subjective expected utility function

$$\lambda^S \mapsto U^S(\lambda^S) \equiv \sum\nolimits_{s \in S} p(s; \epsilon) \sum\nolimits_{y \in Y} \lambda_s(y) \, v(y) \in \mathbb{R}^\infty(\epsilon) \tag{8}$$

on the domain $\Delta(Y^S; \mathbb{R}^\infty_+(\epsilon))$ of $\mathbb{R}^\infty_+(\epsilon)$-valued lotteries $\lambda^S \in \Delta(Y^S; \mathbb{R}^\infty_+(\epsilon))$ whose marginal distributions satisfy $\lambda_s \in \Delta(Y; \mathbb{R}^\infty_+(\epsilon))$ for all $s \in S$. Specifically,

$$\lambda^S \succsim^S \mu^S \iff U^S(\lambda^S) \geq_L U^S(\mu^S)$$

3.7 Lexicographic Expected Utility

The subjective probability $p(s;\epsilon) \in \mathbb{R}_+^\infty(\epsilon)$ of every state $s \in S$ can be expressed as the power series $\sum_{k=0}^\infty p_k(s)\,\epsilon^k$. Thus, the SEU expression (8) can be re-written as the power series $U^S(\lambda^S) \equiv \sum_{k=0}^\infty u_k^S(\lambda^S)\,\epsilon^k$ whose coefficients of successive powers of ϵ are

$$u_k^S(\lambda^S) := \sum\nolimits_{s\in S} p_k(s) \sum\nolimits_{y\in Y} \lambda_s(y)\,v(y) \quad (k = 0, 1, 2, \ldots) \tag{9}$$

But then $\lambda^S \succsim^S \mu^S$, or equivalently $U^S(\lambda^S) \geq U^S(\mu^S)$, if and only if the two respective associated infinite hierarchies of coefficients $\langle u_k^S(\lambda^S)\rangle_{k=0}^\infty$ and $\langle u_k^S(\mu^S)\rangle_{k=0}^\infty$ in the power series satisfy

$$\langle u_k^S(\lambda^S)\rangle_{k=0}^\infty \geq_L \langle u_k^S(\mu^S)\rangle_{k=0}^\infty \tag{10}$$

w.r.t. the usual lexicographic total ordering $\geq_L$ on the space $\mathbb{R}^\infty$ of infinite sequences in $\mathbb{R}$. In this sense, the preference ordering $\succsim^S$ has a lexicographic expected utility representation.

4 Black Swans

4.1 Background

In 82 AD Juvenal (in Satires, VI, 165) had written *"rara avis in terris nigroque simillima cygno"* (a rare bird upon earth, and exceedingly like a black swan). That, however, was merely imaginative irony. Real black swans belonging to the biological species *Cygnus atratus* remained unknown to most of the world before 1697 when Willem de Vlamingh voyaged to what has since become Western Australia. There he became the first European to record seeing living black swans in their native habitat, which included the river he named "Swarte Swaene-Revier" (black swan river). This is now Swan River, which is the main waterway running through the capital city Perth.

Later John Stuart Mill, paraphrasing David Hume, wrote:

> "No amount of observations of white swans can allow the inference that all swans are white, but the observation of a single black swan is sufficient to refute that conclusion."

In elementary philosophy, the existence of black swans has become a classical example of the limits to inferential reasoning.

Taleb's (2007) book provides many vivid examples of events, often related to finance or economics, which he sees as meeting his characterisation of a "Black Swan" event as an "outlier" with "an extreme impact" for which "human nature makes us concoct explanations after the event". The book was written before the recent crisis in global financial markets. Nevertheless, it does discuss several earlier ones like the stock market crash of October 1987 that are often plausibly blamed on faulty statistical models.

Indeed, at an early stage of his book, Taleb defines a "special case of 'gray' swans" which are rare but expected. More precisely, they have probability distributions described by "Mandelbrotian randomness", a particular class of fat-tailed probability distribution following a power law. These distributions put so much weight on outliers, or extreme values, of a random variable $v \in \mathbb{R}$ that, for large enough $k \in \mathbb{N}$, the expectation of the kth power of v, otherwise known as the kth moment of the distribution, becomes infinite. This is in stark contrast to the normal or Gaussian distribution, for which the tail of the distribution is so "thin" that all moments exist.

Yet the main issue with the random value of an asset, especially a derivative security, is typically not whether its distribution has fat or thin tails. Rather, for such assets there is typically a positive probability of losing everything. This potential loss cannot be captured by a Gaussian distribution, or by any "smooth" alternative such as a power law. But there is little really new here, since statisticians and financial economists, along with decision and game theorists, have long been coming to terms with probability distributions which do not correspond to a smooth density function.

4.2 Black Swan Events

Much more challenging than Taleb's "gray swans", however, are the true Black Swans which effectively break our existing scientific models. Indeed, the indisputable existence of the (black) swan species now called *Cygnus atratus* broke all previous biological models of the genus *Cygnus*. While Taleb does recognise that such events could occur, he regards them as "totally intractable", scientifically speaking. Nevertheless, biologists have formulated statistical models intended to forecast probabilistically the likely number of new species that one might expect to find in a poorly explored habitat. And of course economists have developed many models of economic growth with technical progress, which may be approximately treated as the accumulation of many small but typically favourable surprises. A notable example is Schumpeter's (1911, 1934) *The Theory of Economic Development* which sets out the view that, as entrepreneurs innovate, a capitalist market economy is subjected to repeated shocks that cannot be modelled in advance.

More generally, any practical model, especially in the social sciences, must have bounded scope and so must ignore some possibilities. As the statistician George Box wrote: "Essentially, all models are wrong, but some are useful." Should any unmodelled possibility such as a bank run or bank failure occur and have a noticeable impact, it will have to be recognised as an "aberrant" event which, by definition, lies outside the current model.

This is not to deny that any aberrant event could have appeared in an enriched version of the agent's model, if it had been imagined soon enough and then deemed worth modelling. But it was not. Instead, its occurrence demonstrates that the original model is broken and needs modifying accordingly. Such aberrant events lying outside the current model should be distinguished from events within the model which, like Taleb's "gray swans", have extremely low or even zero probability. By contrast, black swan events, unlike those described in Taleb's book, may not even be imagined *ex ante*. Thus, aberrance may be due to a failure of the imagination in constructing a decision model. This may be related to Shackle's (1953) concept of "surprise" — see also Hammond (2007). Indeed, there may be more phenomena in economics that can be explained by "asymmetric imagination" than by the widely used notion of asymmetric information. And not only in economics, but in culture, business, etc.

To summarize, sometimes models may change as their originators anticipate events that had to be excluded originally. To adapt the widely quoted saying by the statistician George Box: "Essentially, all useful models are incompletely specified." The excluded events would become aberrant if they were to occur before they could be included in a more accurate statistical model. Even so, their possible effects on the consequences of modelled current decisions can be allowed for, at least in principle, within a suitable EU decision model allowing an "enlivened" version of the usual decision tree. This is our next topic.

4.3 An Initial Simple Tree

Let Y be a fixed consequence domain. Consider a decision maker whose objective is to maximize the expected value of a von Neumann–Morgenstern utility function (NMUF) $v : Y \to \mathbb{R}$.

Consider an initial (dead) decision tree T:

- with an initial (decision) node n_0,
- at which the *agent* chooses a chance node n_1 in the set $N_1 := N_{+1}(n_0)$ of all nodes that immediately succeed n_0,
- at each of which *chance* determines an immediately succeeding terminal node n_2 in the set $N_2(n_1) := N_{+1}(n_1)$ of all nodes that immediately succeed n_1, using known transition *probabilities* $\pi(n_2|n_1)$ satisfying $\pi(\cdot|n_1) \in \Delta(N_2(n_1))$,
- each of which has a known *final consequence* $\gamma(n_2) \in Y$.

4.4 Initial Evaluation

In this initial simple tree there is a known consequence $\gamma(n_2) \in Y$, of reaching any terminal node n_2. The *initial evaluation* of reaching this node is evidently $w_2(n_2) = v(\gamma(n_2))$.

Working backwards, as usual in dynamic programming, the *conditional expected utility* of reaching any chance node $n_1 \in N_1$ is

$$w_1(n_1) = \mathbb{E}[w_2(n_2)|n_1] = \sum\nolimits_{n_2 \in N_2(n_1)} \pi(n_2|n_1)\, w_2(n_2) \qquad (11)$$

Then an *optimal decision* $n_1^* \in N_1$ is any that maximizes $w_1(n_1)$ with respect to n_1, subject to $n_1 \in N_1$.

The above simple argument is a trivial application to an orthodox "unenlivened" decision model of the *optimality principle* of stochastic dynamic programming. That is, any current decision should be given a *continuation value* equal to the highest possible expected utility resulting from an appropriate plan for all subsequent decisions. Optimality requires the current decision to maximize the expectation of this continuation value.

4.5 Enriched Subtrees

One possible enrichment of the agent's decision model involves a new NMUF $v^+ : Y^+ \to \mathbb{R}$ defined on an enriched model consequence domain $Y^+ \supseteq Y$. But many other enrichments are also possible.

Before we discuss these, note first that the agent can hardly make an unmodelled decision. Accordingly, assume that a necessary and sufficient condition for being able to choose any $n_1 \in N_1$ is that node n_1 is included in the model. Hence the set $N_{+1}(n_0)$ remains fixed. So we assume that any enrichment of the tree takes place only after a particular chosen decision node $n_1^i \in N_{+1}(n_0)$ has already been reached.

What matters, however, is not just how the continuation subtree $T(n_1^i)$ after this particular node is enriched. Also relevant are the potential enrichments of the continuation subtrees $T(n_1)$ at all the other nodes $n_1 \in N_1 \setminus \{n_1^i\}$, since all these possible enrichments ultimately affect the relative expected values of moving to different nodes $n_1 \in N_1$.

Now, starting at each $n_1 \in N_1$, the original continuation subtree $T(n_1)$ had nodes $n_2 \in N_2(n_1)$. Instead there is now an *enriched* continuation subtree $T^+(n_1)$ with:

- an *expanded* set $N_{+1}^+(n_1) = N_2^+(n_1) \supseteq N_2(n_1)$ of immediately succeeding terminal nodes;
- *revised* transition probabilities $\pi^+(n_2^+ | n_1)$ for all $n_2^+ \in N_2^+(n_1)$;
- *revised* consequences $\gamma^+(n_2^+) \in Y^+$ for all $n_2^+ \in N_2^+(n_1)$ with utilities $w_2^+(n_2^+) := v^+(\gamma^+(n_2^+))$.

Instead of (11), the *revised* expected utility of any decision at node n_0 to move to any node $n_1 \in N_1 = N_{+1}(n_0)$ is therefore

$$w_1^+(n_1) := \mathbb{E}^+[w_2^+(n_2^+)|n_1] := \sum\nolimits_{n_2^+ \in N_{+1}^+(n_1)} \pi^+(n_2^+|n_1)\, w_2^+(n_2^+) \tag{12}$$

4.6 Retrospective Evaluation in the Enlivened Tree

In this simple two-stage model, the *enriched tree* T^+ is the extension of T obtained by replacing each continuation subtree $T(n_1)$ $(n_1 \in N_1)$ with its enrichment $T^+(n_1)$. We define the *enlivened tree* as the pair (T, T^+). Unlike botanical tree rings, this includes a complete record of how the tree has grown between:

1. the first period, when it was T;
2. the second period, when it has become T^+.

It is also a mathematical rather than a botanical growth process! For one thing, botanical trees may lose branches in windy conditions, whereas enlivened trees can only expand with time.

Analysed *ex post*, the appropriate decision at initial node n_0 would have been to maximize $w_1^+(n_1^i)$ with respect to $i \in I$. But *ex ante*, only the details of the original model can be used, by definition. What the agent can still do *ex ante*, however, is to recognize that the original evaluation function $w_1(n_1^i)$ may be revised to an as yet unknown and uncertain retrospective evaluation function $w_1^+(n_1^i)$ that ranges over a function space of possible evaluation functions. This is similar in spirit to the work of Koopmans (1964) and Kreps (1992) that allows uncertainty about future preferences — see also Dekel *et al.* (2001, 2007).

In other words, somewhat like Hansen and Sargent (2008, 2011), we can apply a robust decision analysis and choose the initial decision $i \in I$ in order to maximize $\mathbb{E} w_1^+(n_1^i)$ after allowing for uncertainty about the appropriate form of the function $i \mapsto w_1^+(n_1^i)$.

4.7 Cardinally Equivalent Evaluation Functions

Two evaluation functions $w_1, \tilde{w}_1 : N_1 \to \mathbb{R}$ are *cardinally equivalent*, with $w_1 \sim \tilde{w}_1$, just in case there exist:

- an additive constant $\alpha \in \mathbb{R}$
- a positive multiplicative constant $\rho \in \mathbb{R}_+$

such that $\tilde{w}_1(n_1^i) \equiv \alpha + \rho w_1(n_1^i)$.

The *value state space* Ω is defined as the set

- of all non-constant functions $n_1 \mapsto \omega(n_1)$ *normalized* to satisfy

$$\min_{n_1 \in N_1} \omega(n_1) = 0 \quad \text{and} \quad \max_{n_1 \in N_1} \omega(n_1) = 1$$

- together with the *normalized constant function* satisfying $\omega(n_1) = 0$ for all $n_1 \in N_{+1}(n_0)$, which represents complete indifference.

4.8 Uncertain Retrospective Evaluation

Enlivenment replaces the original evaluation function w_1 in T by an *uncertain* retrospective evaluation function w_1^+ derived in the tree T^+, which cannot even be modelled *ex ante*. Because the set N_1 is assumed to be finite, the function $w_1^+ : N_1 \to \mathbb{R}$ ranges over the space $\Omega \subseteq [0,1]^{N_1} \subset \mathbb{R}^{N_1}$ — i.e., Ω is a subset of the unit hypercube in Euclidean space.

4.9 State-Dependent Consequence Domains

In this setting, applying standard subjective probability theory faces an obstacle. The relevant consequences are pairs $(n_1, \omega) \in N_1 \times \Omega$. So the consequence domain $N_1 \times \{\omega\}$ depends on the state $\omega \in \Omega$. This rules out Savage's *constant acts* $a : \Omega \to N_1$ with $a(\omega) = \bar{a}$ for all $\omega \in \Omega$.

In normative decision theory, Hammond (1998b, 1999) suggests a remedy for this kind of state-dependent consequence domain. It is to postulate the existence of an extended NMUF $U : N_1 \times \Omega \to \mathbb{R}$ whose expected value represents preferences $\succsim$ on $\Delta(N_1 \times \Omega)$, when one can choose, in addition to different nodes $n_1 \in N_1$, the probabilities of different states $\omega \in \Omega$.

Given any fixed state $\omega \in \Omega$, the expected values w.r.t. any $\nu \in \Delta(N_1)$ of the two functions $n_1 \mapsto U(n_1, \omega)$ and $n_1 \mapsto \omega(n_1)$ should represent preferences over corresponding lotteries $\nu \in \Delta(N_1)$ and $\nu \times \delta_\omega$. So the two functions $n_1 \mapsto U(n_1, \omega)$ and $n_1 \mapsto \omega(n_1)$ must be cardinally equivalent, for each fixed ω. That is, there must exist mappings $\omega \mapsto \alpha(\omega) \in \mathbb{R}$ and $\omega \mapsto \rho(\omega) \in \mathbb{R}_+$ such that $U(n_1, \omega) \equiv \alpha(\omega) + \rho(\omega)\, \omega(n_1)$.

4.10 Subjective Expected Evaluation

The agent's subjective expected utility objective in the enlivened tree (T, T^+) can (and should) use a *subjective probability measure* P over the Borel subsets of Ω. Then preferences over objective "roulette" lotteries $\nu \in \Delta(N_1)$ are ultimately represented by the *objectively expected* value $\mathbb{E}_\nu V$ of the *subjective expectation* function $N_1 \ni n_1 \mapsto V(n_1)$ defined by

$$V(n_1) := \int_\Omega U(n_1, \omega) P(d\omega) = \int_\Omega [\alpha(\omega) + \rho(\omega)\, \omega(n_1)] P(d\omega) \qquad (13)$$

There is an obvious analogy here with Aumann and Anscombe (1963), who allow combinations of roulette and horse lotteries. An axiomatic justification, however, has yet to be developed, though it should be possible by combining the ideas of Myerson (1979), Fishburn (1982), and Hammond (1998b, 1999).

4.11 Hubris versus Enlivenment

Tractable models are necessarily bounded in scope. Actions may have consequences that are not only unintended, but quite possibly unimagined, and certainly not included in whatever bounded model was used to analyse the agent's decision.

An agent's decision model, like any competent engineer's plan, will typically need to change as and when surprise events outside the model compel attention. Orthodox decision models ignore completely any possibility of model revision. In this sense, they are inherently *hubristic*.

4.12 Could There Be a Metamodel?

A decision model in discrete time amounts to a controlled stochastic process, or equivalently a decision tree that combines chance nodes with decision nodes where the decision is controlled by the decision-maker. Recognizing that the appropriate decision model is itself subject to uncertainty, is it possible, or even desirable, to construct a "metamodel" that embraces all possible decision models?

We will actually consider a simpler question: whether one can or should construct a metamodel in the form of a stochastic "metaprocess" defined on the space of all possible stochastic process models? The result would be a sequence of stochastic processes in which the state space is continually being enriched *unpredictably.*

Now, recall that the stochastic process model is based on Kolmogorov's extension theorem in probability theory. This result states that any "consistent" family of probability laws on finite Cartesian subproducts of an arbitrary collection of component measurable spaces can be extended to a probability law on the whole Cartesian product. The theorem, however, depends on significant topological assumptions such as the existence in each component measurable space of a *compact class* $\mathcal{C}$ of measurable sets — i.e., every sequence of sets in $\mathcal{C}$ whose finite intersections are non-empty has a non-empty infinite intersection — such that the probability of any measurable set must equal the supremum of the probabilities of all its subsets that lie in $\mathcal{C}$.[1] It seems difficult to find a suitable topology on the class of all potentially relevant sequences of stochastic process models which allows an interesting probability measure to exist.

4.13 Should We Look for a Meta Stochastic Process?

El Aleph is a short story published by the distinguished Argentinian author Jorge Luis Borges in 1945. It begins with a quotation from Shakespeare's *Hamlet* Act II, Scene 2

> O God! I could be bounded in a nutshell,
> and count myself a King of infinite space ...

This could be regarded as Shakespeare's poetic description of a key requirement for a metamodel. Eventually we move to the heart of Borges' wonderful story:[2]

> He explained that an Aleph is one of the points in space that contains all other points. ... The Aleph's diameter was probably little more than an inch, but all space was there, actual and undiminished. Each thing (a mirror's face, let us say) was infinite things, since I distinctly saw it from every angle of the universe.

Shortly thereafter the story takes a rather disturbing turn:

> I saw the Aleph from every point and angle, and in the earth the Aleph, and in the Aleph the earth; I saw my own face and my own bowels; I saw your face; and I felt dizzy and wept, for my eyes had seen that secret and conjectured object whose name is common to all men but which no man has looked upon — the unimaginable universe.
>
> I felt infinite wonder, infinite pity.

But eventually something like normality returns:

[1] See Neveu's (1965, p. 82) significant generalization of Kolmogorov's extension theorem, as described in Aliprantis and Border (1999, Section 14.6).

[2] The following brief extracts from `http://www.phinnweb.org/links/literature/borges/aleph.html`, which reproduces the English translation on which Norman Thomas Di Giovanni collaborated with Borges himself.

> Out on the street, going down the stairways inside Constitution Station, riding the subway, every one of the faces seemed familiar to me. I was afraid that not a single thing on earth would ever again surprise me; I was afraid I would never again be free of all I had seen. Happily, after a few sleepless nights, I was visited once more by oblivion.

A later postscript includes some explanation for Borges' choice of title:

> As is well known, the Aleph is the first letter of the Hebrew alphabet. Its use for the strange sphere in my story may not be accidental. For the Kabbala, the letter stands for the *En Soph*, the pure and boundless godhead; it is also said that it takes the shape of a man pointing to both heaven and earth, in order to show that the lower world is the map and mirror of the higher; for Cantor's *Mengenlehre* [set theory], it is the symbol of transfinite numbers, of which any part is as great as the whole.

Perhaps the moral of Borges' story is that in the end we should be relieved about how mathematically and conceptually intractible the problem of finding a stochastic metaprocess appears to be.

5 Concluding Remarks

A descriptive decision theory stands or falls by its capacity to explain what we observe. A prescriptive decision theory, on the other hand, stands or falls by its capacity to offer a normatively appealing approach to decision making. This work has set out alternative departures from standard prescriptive decision theory. These departures have been designed to deal separately with the three key phenomena of catastrophic risk, rare events, and true black swan events that transcend whatever decision model we may currently be using.

The work by Chichilnisky (1996, 2000, 2009, 2010) has set out heroically to deal with all these three phenomena within one integrated framework. In doing so, however, she follows Savage (1954) in relaxing the usual countable additivity property of probability measures, thus allowing probabilities that are only finitely additive. A conjecture to be settled by future research is that the same three phenomena could be accommodated within a different integrated framework which retains a countably additive probability measure. This framework would allow:

1. the kind of distinction between catastrophic and non-catastrophic consequences that was introduced in Section 2;

2. for rare events, non-Archimedean probabilities of the kind discussed in Section 3, but extended from a finite sample space S to a general measurable space $(S, \mathcal{S})$;

3. for true black swan events, enlivened trees of the kind sketched briefly in Section 4, with preferences represented by subjective expected utility based on extended probability measures over states of the world that correspond to possible retrospective evaluation functions defined for every modelled decision.

Note finally that rationality within bounded decision trees allows a *restricted* revealed preference hypothesis, applying only to options that receive serious consideration. But decision trees almost inevitably become *enlivened* in case the decision maker is forced to recognize the possibility of events which were excluded from earlier decision models. These unmodelled events are *truly unknown* "black swans", like the species *cygnus atratus* was to Europeans before Dutch explorers reached Western Australia. Such unmodelled events are *completely different* from the "highly improbable" but modelled events referred to as "grey swans" in Taleb (2007). Indeed, Taleb dismisses true black swans as completely intractable.

References

[1] Aliprantis, C.D. and K.C. Border (1994, 2nd. edn. 1999) *Infinite Dimensional Analysis: A Hitchhiker's Guide* (Berlin: Springer).

[2] Anscombe, F. and R.J. Aumann (1963) "A Definition of Subjective Probability" *Annals of Mathematical Statistics* 34: 199–205.

[3] Arrow, K.J. (1965) *Aspects of the Theory of Risk-Bearing* (Helsinki: Yrjö Jahnssonin Säätiö).

[4] Balasko, Y. (1978) "Economic Equilibrium and Catastrophe Theory: An Introduction" *Econometrica* 46: 557–569.

[5] Chichilnisky, G. (1996) "Updating von Neumann Morgenstern Axioms for Choice under Uncertainty" In: *Proceedings of a Conference on Catastrophic Risks* (Toronto: The Fields Institute for Mathematical Sciences).

[6] Chichilnisky, G. (2000) "An Axiomatic Approach to Choice under Uncertainty with Catastrophic Risks" *Resource and Energy Economics* 22: 221–231.

[7] Chichilnisky, G. (2009). "The Topology of Fear" *Journal of Mathematical Economics* 45: 807–816.

[8] Chichilnisky, G. (2010) "The Foundations of Statistics with Black Swans" *Mathematical Social Sciences* 59: 184–192.

[9] Dekel, E., B.L. Lipman, and A. Rustichini (2001) "Representing Preferences with a Unique Subjective State Space," *Econometrica* 69: 891–934.

[10] — with T. Sarver (2007) "— : Corrigendum" *Econometrica* 75: 591–600.

[11] Drèze, J.H. (1962) "L'utilité sociale d'une vie humaine" *Revue Française de Recherche Opérationnelle*, 23: 93–118.

[12] Fishburn, P.C. (1982) *The Foundations of Expected Utility* (Dordrecht: D. Reidel).

[13] Halpern, J.Y. (2009) "A Nonstandard Characterization of Sequential Equilibrium, Perfect Equilibrium, and Proper Equilibrium" *International Journal of Game Theory* 38: 37–49.

[14] Halpern, J.Y. (2010) "Lexicographic Probability, Conditional Probability, and Nonstandard Probability" *Games and Economic Behavior* 68: 155–179.

[15] Hammond, P.J. (1994) "Elementary Non-Archimedean Representations of Probability for Decision Theory and Games" In P. Humphreys (ed.) *Patrick Suppes: Scientific Philosopher, Vol. I: Probability and Probabilistic Causality* (Kluwer Academic Publishers, 1994), ch. 2, pp. 25–59.

[16] Hammond, P.J. (1997, 1999) "Non-Archimedean Subjective Probabilities in Decision Theory and Games" Stanford University Dept. of Econ. Working Paper No. 97-038; abbreviated version in *Mathematical Social Sciences* 38: 139–156.

[17] Hammond, P.J. (1998a) "Objective Expected Utility: A Consequentialist Perspective" In Barberà, S., Hammond, P.J., and Seidl, C. (eds.) *Handbook of Utility Theory, Vol. 1: Principles* (Boston: Kluwer Academic Publishers) ch. 5, pp. 145–211.

[18] Hammond, P.J. (1998b) "Subjective Expected Utility" In Barberà, S., Hammond, P.J., and Seidl, C. (eds.) *Handbook of Utility Theory, Vol. 1: Principles* (Boston: Kluwer Academic Publishers) ch. 6, pp. 213–271.

[19] Hammond, P.J. (1999) "Subjectively Expected State-Independent Utility on State-Dependent Consequence Domains" in M.J. Machina and B. Munier (eds.) *Beliefs, Interactions, and Preferences in Decision Making* (Dordrecht: Kluwer Academic), pp. 7–21.

[20] Hammond, P.J. (2007) "Schumpeterian Innovation in Modelling Decisions, Games, and Economic Behaviour" *History of Economic Ideas* XV: 179–195.

[21] Hammond, P.J. (2009) "Adapting to the entirely unpredictable: black swans, fat tails, aberrant events, and hubristic models" `http://www2.warwick.ac.uk/fac/soc/economics/research/centres/eri/bulletin/2009-10-1/hammond/`.

[22] Hansen, L.P and T.J. Sargent (2008) *Robustness* (Princeton University Press: Princeton, NJ).

[23] Hansen, L.P and T.J. Sargent (2011) "Wanting Robustness in Macroeconomics", in B.M. Friedman and M. Woodford (eds.), *Handbook of Monetary Economics, Volume 3B* (Elsevier) ch. 20, pp. 1097–1157.

[24] Jensen, N.E. (1967) "An Introduction to Bernoullian Utility Theory, I: Utility Functions" *Swedish Journal of Economics* 69: 163–183.

[25] Jones-Lee, M.W. (1974) "The Value of Changes in the Probability of Death or Injury" *Journal of Political Economy* 82: 835–849.

[26] Kolmogorov, A.N. (1933, 1956) *Grundbegriffe der Wahrscheinlichkeitsrechnung* (Berlin: Springer); translated as *Foundations of Probability* (New York: Chelsea).

[27] Koopmans, T.C. (1964) "On Flexibility of Future Preference" in M.W. Shelly and G.L. Bryan (eds.) *Human Judgments and Optimality* (New York: John Wiley), ch. 13, pp. 243–254.

[28] Kreps, D.M. (1992) "Static Choice in the Presence of Unforeseen Contingencies," in P. Dasgupta, D. Gale, O. Hart, and E. Maskin (eds.), *Economic Analysis of Markets and Games: Essays in Honor of Frank Hahn* (Cambridge, Mass.: M.I.T. Press), pp. 258–281.

[29] Lightstone, A.H. and A. Robinson (1975) *Nonarchimedean Fields and Asymptotic Expansions* (Amsterdam: North-Holland).

[30] McGonagall, W. (1880) "The Tay Bridge Disaster" `http://www.mcgonagall-online.org.uk/gems/the-tay-bridge-disaster`.

[31] Myerson, R.B. (1979) "An Axiomatic Derivation of Subjective Probability, Utility, and Evaluation Functions" *Theory and Decision* 11: 339–352.

[32] Neveu, J. (1965) *Mathematical Foundations of the Calculus of Probability* (San Francisco: Holden-Day)

[33] Rényi, A. (1955) "On a New Axiomatic Theory of Probability" *Acta Mathematica Academiae Scientiarum Hungaricae* 6: 285–335.

[34] Rényi, A. (1956) "On Conditional Probability Spaces Generated by a Dimensionally Ordered Set of Measures" *Theory of Probability and its Applications* 1: 61–71.

[35] Robinson, A. (1973) "Function Theory on Some Nonarchimedean Fields" *American Mathematical Monthly: Papers in the Foundations of Mathematics* 80: S87–S109.

[36] Savage, L.J. (1954, 1972) *The Foundations of Statistics* (New York: John Wiley; and New York: Dover Publications).

[37] Schumpeter, J.A. (1911; 2nd edn. 1926) *Theorie der wirtschaftlichen Entwicklung; Eine Untersuchung über Unternehmergewinn, Kapital, Kredit, Zins und den Konjunkturzyklus* (München und Leipzig: Duncker und Humblot).

[38] Schumpeter, J.A. (1934, 1961) *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle*. Translated from the German by Redvers Opie; with a new Introduction by John E. Elliott.

[39] Selten, R. (1975) "Re-examination of the Perfectness Concept for Equilibrium Points of Extensive Games" *International Journal of Game Theory* 4: 25–55.

[40] Shackle, G.L.S. (1953) "The Logic of Surprise" *Economica, New Series* 20: 112–117.

[41] Taleb, N.N. (2007) *The Black Swan: The Impact of the Highly Improbable* (New York: Random House).

[42] Thom, R. (1973) *Stabilité Struturelle et Morphogéneèse* (Paris: Interéditions); translated (1976) as *Structural Stability and Morphogenesis* (Reading, MA: W.A. Benjamin).

[43] Villegas, C. (1964) "On Qualitative Probability σ-Algebras," *Annals of Mathematical Statistics* 35: 1787–1796.

[44] von Neumann, J. and O. Morgenstern (1944; 3rd edn. 1953) *Theory of Games and Economic Behavior* (Princeton: Princeton University Press).

[45] Zeeman, E.C. (1976) "Catastrophe Theory" *Scientific American* (April), pp. 65–70 and 75–83.