

DECISION-MAKING UNDER UNCERTAINTY

Probability and Forecasting

Tetlock · Silver · Duke · Taleb

PAPER 03 / 03

A Furtherance reference compendium on the intellectual traditions behind organisational design, leadership, execution, and culture.

IN THIS PAPER

The series closes where uncertainty bites hardest: forecasting the future and deciding under it. This paper covers the evidence on forecasting skill, the discipline of probabilistic thinking, decision-making as betting, and the limits of prediction in a fat-tailed world. The question: when is the future forecastable – and what do we do when it is not?

The four voices divide into a method and a warning. Tetlock shows forecasting is a real, measurable skill; Silver gives it a probabilistic discipline; Duke turns it into a way of deciding. Taleb then draws the boundary: in domains ruled by rare, extreme events, prediction fails, and the task becomes robustness.

Philip Tetlock *Superforecasting, 2015*

Forecasting research · the Good Judgment Project

CORE THESIS

Expert forecasts are, on average, little better than chance – but forecasting is a **learnable skill**, and a minority of **superforecasters** reliably beat the rest. The decisive trait is cognitive style: **foxes**, who hold many small ideas and update often, out-predict **hedgehogs**, who force events through one big theory.

FRAMEWORK

Trait of the superforecaster	In practice
Probabilistic thinking	Answers in degrees, not yes/no
Bayesian updating	Revises in small steps as evidence arrives
Decomposition	Breaks a question into tractable parts
Actively open-minded	Seeks disconfirming views; treats beliefs as testable

Accuracy is scored with **Brier scores**, making forecasting accountable rather than rhetorical – the fox-hedgehog distinction is measured, not asserted.

METHODOLOGY

Large, long-running forecasting tournaments – Expert Political Judgment and the IARPA-sponsored Good Judgment Project – scoring thousands of forecasters over years.

SIGNIFICANCE AND CRITIQUE

A rare, rigorous evidence base showing that forecasting skill exists and can be trained and scored. The critique: the questions are mostly geopolitical and resolve on short horizons, and the findings may not extend to unique or genuinely unprecedented events.

Nate Silver *The Signal and the Noise, 2012*

Applied statistics · probabilistic forecasting

CORE THESIS

Good forecasting means separating **signal** from **noise** and thinking in calibrated **probabilities** rather than point predictions. Most forecasts fail from overconfidence and **overfitting** – mistaking noise for pattern – and the remedy is disciplined **Bayesian** updating.

FRAMEWORK

Principle	What it demands
Signal vs noise	Distinguish the real pattern from random fluctuation
Bayesian updating	Start from a prior; revise with each new piece of evidence
Calibration	A 70% forecast should happen about 70% of the time
Avoiding overfitting	Prefer simple models that generalise over complex ones that flatter the past

METHODOLOGY

Applied statistical forecasting across domains – elections, sport, weather, economics – written for a general audience.

SIGNIFICANCE AND CRITIQUE

Popularised probabilistic, Bayesian thinking and the language of calibration for a wide readership. The critique: it is applied synthesis rather than original scholarship, and high-profile probabilistic forecasts are easily misread by audiences who want certainty.

Annie Duke *Thinking in Bets*, 2018

Decision science · poker and applied judgement

CORE THESIS

Every decision is a **bet** on an uncertain future, so decision quality must be judged separately from outcome quality. Judging a decision by how it turned out – “**resulting**” – is the central error: good decisions can yield bad outcomes, and bad decisions good ones, over any short run.

FRAMEWORK

Idea	What it means
Resulting	The error of grading a decision by its outcome alone
Thinking in bets	Express beliefs as probabilities and stakes
Truthseeking groups	Use a candid pod to check reasoning, not to confirm it
Kill criteria	Decide in advance what would make you quit

METHODOLOGY

A practitioner synthesis drawing on a career in professional poker and on cognitive science – decision-making where probability, incomplete information and stakes are explicit.

SIGNIFICANCE AND CRITIQUE

A sharp, memorable translation of probabilistic thinking into everyday decision practice; “resulting” and “kill criteria” are genuinely useful tools. The critique: it is practitioner and popular in register, offering translation rather than new theory.

Nassim Nicholas Taleb The Black Swan, 2007

Probability and risk · the Incerto

CORE THESIS

The most consequential events are **black swans** – rare, high-impact, and rationalised only in hindsight – and they are essentially unforecastable. In **Extremistan**, the domain of fat tails, conventional statistics and forecasting mislead dangerously. The response is not better prediction but **robustness**: manage exposure, avoid ruin, and build systems that gain from disorder.

FRAMEWORK

Concept	Meaning
Black swan	A rare, high-impact event, explained away after the fact
Mediocristan vs Extremistan	Thin-tailed, forecastable domains vs fat-tailed, unforecastable ones
Antifragility	Systems that improve under stress and volatility
Barbell strategy	Pair extreme safety with small, capped high-upside bets
Skin in the game	Decision-makers must bear the downside of their choices

METHODOLOGY

Probability theory argued through essay, drawing on a career as a derivatives trader preoccupied with rare, extreme risk.

SIGNIFICANCE AND CRITIQUE

A profound corrective to forecasting overconfidence in fat-tailed domains, and the convexity-and-robustness framing is genuinely useful. The critique: the work is polemical and hard to operationalise, and at times overstates unpredictability – dismissing the forecasting skill the rest of this paper documents in regular domains.

A NOTE ON EVIDENCE

A note on scope: the evidence for trainable forecasting skill comes largely from tournaments of well-defined, short-horizon questions, and does not straightforwardly extend to unique, long-horizon or fat-tailed events – exactly the territory Taleb marks out. Read this paper as domain-dependent: probabilistic forecasting where feedback is good, robustness where it is not. The two are complements, not rivals.

Synthesis

Theorist	Core idea	Key tool	On predictability
Tetlock	Forecasting is a learnable skill	Foxes; Brier scores	Skilled in regular domains
Silver	Think in calibrated probabilities	Bayesian updating	Signal is real but limited
Duke	Decisions are bets	Resulting; kill criteria	Judge process, not outcome
Taleb	Rare events dominate	Antifragility; barbell	Fat tails are unforecastable

FOR THE OPERATOR

The series resolves into a stance on uncertainty. Where the environment is regular and feedback is good, forecasting is a real skill – done probabilistically and Bayesianly (Tetlock, Silver), and judged on process not outcome (Duke). Where tails are fat and feedback is poor, prediction fails and the task shifts to robustness and avoiding ruin (Taleb). For the operator: calibrate your method to the domain. That closes the Decision-Making Under Uncertainty series – from the question of whether we are rational, through the case that we are boundedly adaptive, to the discipline of acting when the future will not be known.