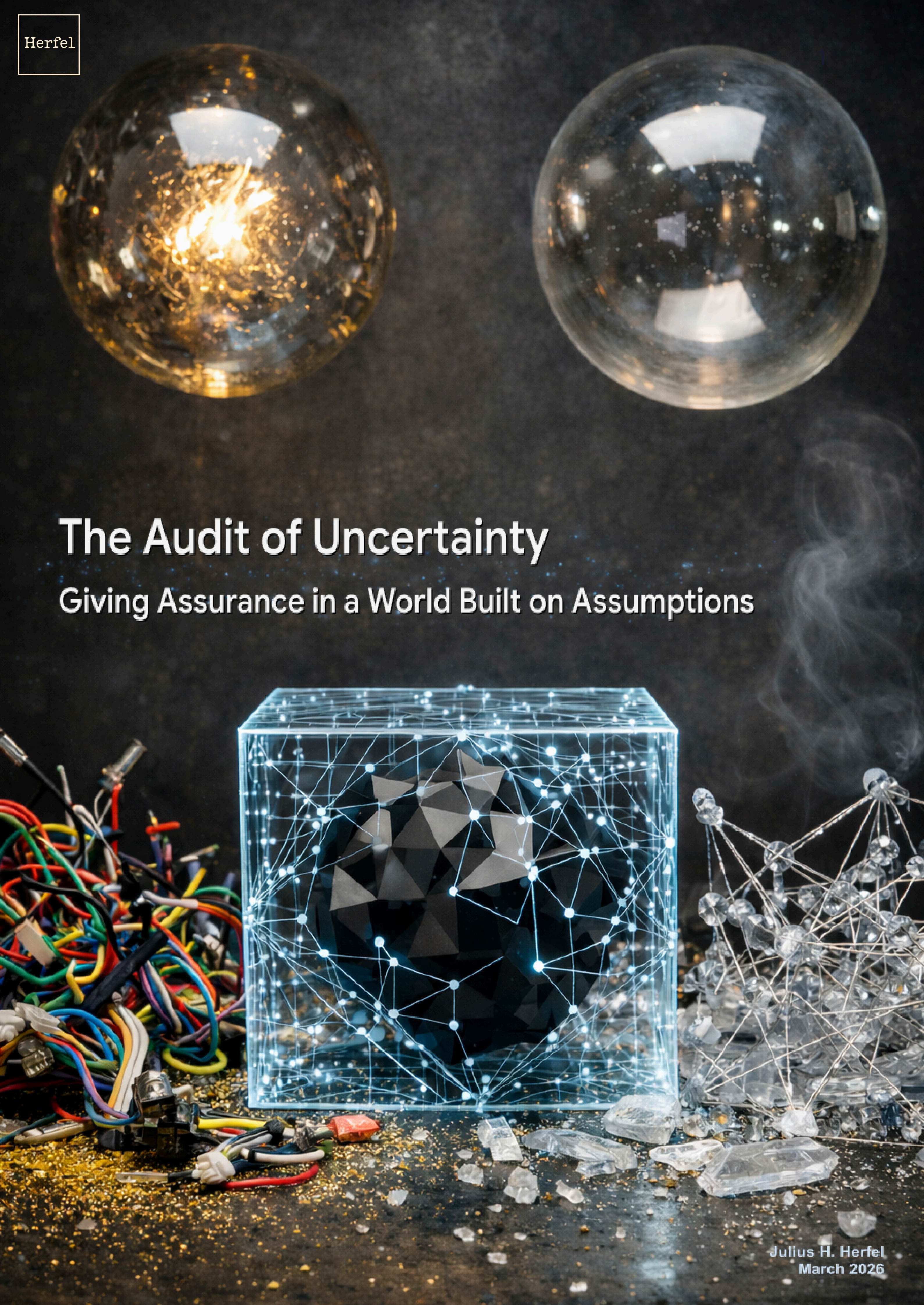
Herfel



The Audit of Uncertainty

Giving Assurance in a World Built on Assumptions

Julius H. Herfel
March 2026

Executive Summary

The profession of Internal Audit has quietly shifted its centre of gravity: from verifying recorded facts to providing assurance over judgements, estimates, forecasts and models. Financial statements and risk management frameworks now fundamentally depend on assumptions about economic conditions, agent behaviour, market structure, and future events. Assumptions are where audit risk hides.

To make assumptions auditable, this whitepaper presents a three-dimensional taxonomy classifying assumptions by status (how entrenched they are), location (where they sit in the modelling and decision-making process), and function (what role they play in the model's logic). This taxonomy enables auditors to systematically identify assumptions, assess their risk profile, select appropriate testing methods, and evaluate governance and process adequacy.

The governing principle is fitness for purpose. Assumptions do not need to be "right" in an absolute sense, but they must be appropriate for their intended use. The auditor's responsibility is to ensure that management can articulate why their chosen assumptions are suitable, that alternatives have been considered, that sensitivity to changes is understood, and that limitations are transparent.

This whitepaper provides three core audit tools - sensitivity analysis, scenario analysis, and backtesting - with structured mini audit programmes for each. It provides more granular examples for two domains: derivative pricing and Level 3 instrument valuation, and operational deposit modelling. Finally, it introduces a concept for AI-assisted assumption identification and classification that can support enterprise-wide assumption management.

Foreword - The Changing Role of Audit

Ticking boxes, looking for typos in procedures, digging through large tables of numbers. That is what many think Audit does. Not quite. Not in 2026. In the early days of Audit as a profession, it focused on the verification of recorded facts. The auditor's role was to confirm that transactions had occurred as documented, and that balances reconciled to underlying records. This paradigm rested on a stable environment where most material items subject to audit were either objectively determinable or governed by unambiguous rules. This changed with the advent of provisions to account for potential loan losses.

The London & General Bank case of 1895 marked an early inflection point. Auditors were held liable for misfeasance because they identified that the security for loans was insufficient but failed to clearly communicate this material fact to shareholders, using instead vague language that concealed the company's actual position. The case established that an auditor's duty extends beyond arithmetical verification to conveying the true financial position, ensuring that shareholders receive "information" rather than merely the "means of information" regarding the sufficiency of asset security. However, the systematic integration of judgement, models, and forward-looking estimates into the core of financial reporting and risk management remained limited for nearly a century thereafter.

From Verification to Assurance over Judgement

The transition accelerated following the end of the Bretton Woods system and the oil crisis of the 1970s, which introduced volatility into exchange rates, interest rates, and commodity prices. Financial institutions began employing quantitative models to measure and manage risks that had previously been addressed qualitatively, if at all. The Basel I framework in 1988 introduced regulatory capital requirements, which Basel II subsequently made risk-sensitive, later also permitting banks to use internal models for capital calculation. Fair value concepts existed in accounting for decades, but the explicit definition of fair value as an exit price was standardised by FAS 157 in 2006 in the US and IFRS 13 in 2011 internationally (effective 2013). IFRS 13 established that fair value should be measured as the price that would be received to sell an asset or paid to transfer a liability in an orderly transaction between market participants at the measurement date. This is straightforward for liquid instruments. For rarely traded or highly bespoke instruments - exotic options or complex structured finance vehicles, for example - this requires numerous assumptions and approximations, including future funding costs and the market's perception of the institution's own and its counterparty's default probability.

The 2008 Global Financial Crisis demonstrated the failure of existing audit and risk management frameworks. The crisis was not primarily caused by fraudulent transactions or bookkeeping errors. It was driven to a large extent by assumptions that proved invalid under stress, by models that failed to capture tail risk and correlation behaviour, and by a collective underestimation of structural uncertainty. Liquidity was assumed to be always available. Default correlations were assumed to remain stable. Housing prices were assumed not to decline simultaneously across geographic regions. These assumptions, embedded in pricing models, risk models, and business strategies, were treated as facts until they collapsed.

Audit's role consequently had to evolve from verification of facts to assurance over the fitness for purpose of judgements, estimates, and models. This shift is not temporary. Nowadays, uncertainty is the rule. Financial statements now contain material balances determined by management judgement and model outputs rather than contractual terms or market prices. Expected credit loss provisioning under IFRS 9, fair value measurement under IFRS 13, actuarial assumptions in pension accounting, and goodwill impairment testing all require estimation of future events and application of expert judgement.

CCAR, ICAAP, and ILAAP are inherently uncertain, just like recovery plans, financial forecasts, M&A models and climate stress tests. Governance, which has received increased focus from regulators and internal auditors since the 1980s, is often based on implicit rules, judgement and opinion; nothing auditors can reliably verify. Instead of looking for something to benchmark against, auditors now have to dig into assumptions underlying judgement.

The Primacy of Assumptions

In this environment, assumptions, not data, are the dominant risk driver. A model's mathematical sophistication is irrelevant if its foundational assumptions are inappropriate. The mathematical structure of a model is conditional on its assumptions. Consider a standard discounted cash flow valuation. The formula is not controversial: the present value of a series of cash flows discounted at an appropriate rate. However, every input is assumption-dependent. What are the expected cash flows? What is the appropriate discount rate? What is the terminal growth rate? What is the probability distribution of outcomes? Different assumptions produce different valuations using the identical formula. Two analysts applying the same model to the same entity can produce materially different fair value estimates based solely on differences in assumptions about growth rates, discount rates, or risk adjustments.

The model will not tell you whether your assumptions make sense. However, the model structure can in itself be an assumption. A Value-at-Risk model can be based on parameters, historical data, or the output of a Monte Carlo simulation. While this is often driven by practicalities and limitations, the use of historical data implicitly assumes the past is a good predictor of the future - a common assumption, often fit for purpose, but also often not accurate.

The use of a model's output also requires critical evaluation. In the case of collateralised debt obligations (CDOs), models built to quantify risk, specifically default probabilities, were used to price CDOs. That an instrument's risk impacts its price is a reasonable pricing consensus, but when risk became the price itself and the risk was underestimated, the crash in the US housing market triggered CDO prices to collapse, resulting in losses larger than the actual losses from defaults of the underlying obligations.

Fitness for Purpose as the Governing Principle

In practice, fitness for purpose is the only defensible principle for assessing assumptions. An assumption may be technically incorrect yet functionally adequate. Special care is required, however, to ensure that incorrect assumptions do not lead to misstatements in financial reporting or inadequate model outputs. Conversely, an assumption may be theoretically sound yet inappropriate for the specific application. The auditor's job is not to substitute management's judgement for theirs, nor to declare assumptions "right" or "wrong," unless supported by thorough evidence, but to ensure management can articulate why their chosen assumptions are suitable for their needs, that alternative assumptions have been considered, and that sensitivity to changes in those assumptions is understood.

This whitepaper provides a structured framework for auditing assumptions across financial models, judgement, and decision systems. It classifies assumptions along three dimensions, presents core audit tools for testing them, applies the framework to two topical examples, derivative pricing and operational deposit modelling, and introduces a concept for AI-assisted assumption management.

Table of Contents

Introduction

Financial reality is not built on facts carved in stone. Even something as basic as the fair value rests on judgement. It is constructed through interconnected layers of belief, each dependent on the integrity of the layer beneath it. The Layered Beliefs Framework (LBF) describes four distinct layers.

The Four Layers

Assumptions

The Assumptions layer contains foundational beliefs about how the world operates: beliefs about economic relationships, agent behaviour, market structure, and the stability or variability of observed phenomena. Assumptions feed into models and decisions. Often they are not derived from data, and when they are, they still determine how we interpret that data.

Models

The Model layer formalises assumptions into quantitative or logical structures. Models translate assumptions into outputs: prices, risk metrics, forecasts, or decisions. A model is a representation, not reality itself. It is useful to the extent that it simplifies without materially distorting.

Consensus

The Consensus layer aggregates individual model outputs and expert judgements to produce a "representative truth." This layer includes agreements - explicit and implicit - on the factors impacting a market price, consensus valuations for unobservable instruments, and collective professional judgements such as accounting estimates or regulatory classifications. Consensus is a diffuse and informal coordination mechanism. It does not guarantee accuracy.

The Consensus layer is the element that differentiates the LBF from most traditional frameworks. It acknowledges that market prices do not reflect some unique "true" equilibrium discovered through supply and demand of millions of neutral agents in an efficient market, but rather a normative stage prior to the market which informs market participants' actions. The Consensus layer does not validate Efficient Market Hypothesis; it highlights that market "efficiency" is an emergent property of shared beliefs, a normative stage where market participants agree on a value before capital is exchanged.

Markets

The Market layer acts upon the consensus. Prices are paid, capital is allocated, positions are taken, and decisions are executed. Where consensus is the agreement of relevant parties on value (the belief), the market is the irrevocable exchange of capital based on that agreement (the action). The market layer converts belief into realised outcomes, which then feed back into the assumptions layer through observed data. Whereas the consensus is reached by actors with sufficient powers, the market result is applicable to all participants.

The Market layer is where beliefs become consequential. A consensus valuation of a CDO tranche is an opinion; a trade at that price is an irreversible commitment of capital. This distinction matters because market actions create observable data points that feed back into assumptions, but the data they generate is shaped by the beliefs that preceded them. Market prices are therefore not independent evidence of "true" value, but the crystallisation of prior consensus, which may itself be flawed.

Connectors, Recursion, and Bias

Each layer depends on what lies beneath it. Flawed assumptions invalidate even sophisticated models; model errors distort consensus; corrupted consensus misallocates capital. The dependency is transmitted through connectors which serve as inputs to every layer: assumptions connect to models through opinions and data; models connect to consensus through uncertainty and meaning; consensus connects to markets through prices and metrics.

The framework loops back on itself. Market outcomes feed back into future assumptions, creating feedback loops that can amplify both accuracy and error. Yet it is a framework. It can explain market realities and many crises of the past, but should not be treated as a law of nature. Treating a model, an assumption, or a framework as a given truth and not considering its uncertainty is always a mistake.

The most fundamental input to the framework, however, is bias. Most cognitive biases are evolutionarily inherited. Our ancestors evolved brains optimised for living in the wild. Mathematics evolved much later; models even after that. Our biases are not calibrated for making good models, but for making decisions fast even if that impedes precision, filtering information even if it distorts truth, and aiming for social cohesion even if that introduces unfairness. As we walk through the types of assumptions, we will always consider the potential biases. Notably, a well-documented bias is the bias blind spot (Pronin et al., 2002), the tendency to recognise the influence of biases on other people's judgements but not our own. Everybody is biased. Structured taxonomies help us recognise this in ourselves.

Why Assumptions Require Structured Classification

Without a structured classification system, audit reviews run the risk of being ad hoc, missing critical assumption types, or applying inappropriate testing methods. A simplifying assumption (one that deliberately reduces the complexity of reality) requires different audit procedures than a speculative assumption (a recognised guess about an uncertain future). An assumption embedded in a model's structural architecture propagates error differently than one about agent behaviour. An assumption that establishes a baseline state has different failure modes than one that defines a causal relationship.

Structured classification enables auditors to identify which assumptions are present in a given model or process, determine which assumptions are material to output, select the right tests for each assumption type, assess whether governance is commensurate with assumption risk, and evaluate whether sensitivities and limitations are understood. The taxonomy presented in this whitepaper is not an academic exercise but a practical tool for systematic assumption identification and evaluation.

This whitepaper does not explicitly differentiate between aleatoric uncertainty (randomness in a system) and epistemic uncertainty (uncertainty from a lack of knowledge). The difference is not benign, as I will argue in an upcoming discussion paper on volatility modelling. Some assumptions we discuss are clearly aleatoric, others clearly epistemic. Many are both. The assumptions taxonomy provided and the audit guidance given capture the differences, without explicitly categorising them each time. It is absolutely possible to extend the framework with these assumption types; however, as they are not mutually exclusive, I found it more appropriate to reflect the difference in the types of assumptions in the taxonomy and propose audit steps.

Part I - Classifying Assumptions

Classification by Status

The status of an assumption describes how entrenched a belief is within the modelling context and the broader professional or market environment. Status determines the level of scrutiny an assumption receives, the governance required, and the difficulty of challenging it.

Simplifying Assumptions

Definition: A simplifying assumption is deliberately adopted to make a complex problem tractable by reducing dimensionality, aggregating heterogeneous elements, or abstracting away details that are believed not to materially affect the output for the intended purpose.

Examples

In determining the allowance for expected credit losses on a retail credit card portfolio, a bank may assume that all borrowers within a given credit score band will exhibit the same probability of default, despite knowing that individual circumstances vary. This aggregation is necessary to model a portfolio of millions of accounts.

A trading desk building a Value-at-Risk (VaR) model for its foreign exchange portfolio may assume that exchange rate returns are normally distributed. The actual distribution has fatter tails, but the normal distribution simplifies calculation and may be adequate for 95% confidence intervals under ordinary market conditions.

Cognitive Biases

Simplifying assumptions are often driven by availability heuristics. Modellers gravitate towards familiar distributions and structures because these are readily available in standard software and educational materials, though analytical tractability and well-understood mathematical properties also play a legitimate role. Anchoring bias also contributes; once a simplifying assumption is adopted in an initial model version, subsequent iterations may retain it without re-examination simply because it was "the original assumption."

Why This Status Matters for Audit

We cannot model without simplifying. Simplifying assumptions aim to achieve on a micro-level what models themselves do on a macro level: break a complex world down into a single, basic narrative. The audit objective is not to eliminate simplifying assumptions but to ensure that the simplification does not introduce material distortion for the model's intended use. A simplification appropriate for one purpose may be inappropriate for another. Assuming normally distributed returns may be adequate for daily VaR at 95% confidence but grossly inadequate for Expected Shortfall at 97.5% confidence, which measures tail risk.

Key Expected Management Controls/ Possible Audit Tests

Sensitivity analysis should vary the simplifying assumption to quantify its impact on outputs. If output changes are within acceptable tolerance, the simplification is defensible. Benchmarking compares the simplified approach against a more complex alternative (e.g. normal distribution VaR against historical simulation VaR) to assess the magnitude of difference. Segmentation testing examines whether sub-segmentation of aggregated populations would materially change results. Documentation review verifies that the assumption is documented, its purpose explained, materiality assessed, and approval obtained from an appropriate authority.

Assumptions Stating the Obvious

Definition: An assumption stating the obvious reflects a widely accepted empirical regularity or market principle. It is treated as self-evident. While it can be wrong under certain conditions, departing from it introduces new risks, and audit needs to carefully assess whether a challenge is warranted.

Examples

An impairment test for a financial asset may assume that risk and return are positively correlated. Higher-risk instruments require higher returns to attract investors. While this assumption can fail under specific conditions (e.g. flight-to-quality scenarios where risk-free assets appreciate), it reflects long-run empirical regularity.

A credit pricing model may assume that obligors with higher credit ratings have lower probabilities of default. This reflects ratings agencies' methodologies and centuries of credit market history. Departing from this assumption without strong evidence would create model risk.

Cognitive Biases

These assumptions often persist due to conformity bias. There is social and professional pressure to align with consensus. Challenging an "obvious" assumption may be perceived as contrarian without purpose. Confirmation bias reinforces them. Practitioners preferentially notice evidence supporting the obvious relationship and discount anomalies as temporary or idiosyncratic.

Why This Status Matters for Audit

Assumptions stating the obvious are low-risk under normal conditions but can become dangerous during regime shifts. Auditors must assess whether management has considered boundary conditions. These are situations under which the assumption might fail. Audit should also ascertain whether management has assessed whether those boundary conditions are relevant to the model's application.

Key Expected Management Controls/ Possible Audit Tests

Boundary analysis documents the conditions under which the assumption holds and those under which it fails, assessing whether the model operates in contexts where failure is plausible. Historical backtesting examines whether the assumption held during past stress periods. Opposite assumption testing performs a thought experiment: what would occur if the opposite were true? If the opposite produces absurd results under all conditions, the original is robust; if the opposite is plausible under certain conditions, those conditions should be monitored. Governance review verifies periodic review and that contradictory evidence is considered.

Sacred Assumptions

Definition: A sacred assumption is normatively protected or accepted as absolute truth within a professional community or market context. Challenging it is perceived as either naive or radical. Sacred assumptions persist not because of evidence but because challenging them is career-limiting.

Examples

Prior to the Global Financial Crisis, the assumption that AAA-rated sovereign debt issued by developed economies was risk-free for collateral and capital purposes was treated as sacred. Regulatory frameworks facilitated this, most notably through the 0% risk weight for AAA sovereigns under the Standardised Approach and the widespread use of national discretions. Challenging it would have required challenging the entire regulatory architecture.

For decades, the assumption that residential mortgage defaults across different US geographic regions were largely uncorrelated was sacred in mortgage-backed securities pricing and rating methodologies. Historical data supported low correlation, and the assumption underpinned the entire securitisation market. Questioning it was uncommon until correlations spiked during the crisis.

Cognitive Biases

Sacred assumptions are sustained by conformity bias and authority bias; they are endorsed by regulators, rating agencies, academic authorities, or industry-wide practice. Groupthink prevents meaningful challenge; dissenting views are suppressed or ignored because they threaten the consensus. Fundamental attribution error may in some scenarios also play a role: when sacred assumptions fail, the failure is attributed to external shocks rather than flaws in the assumption itself.

Why This Status Matters for Audit

Sacred assumptions are the most difficult to identify and challenge. Doing so may place the auditor or the auditee at a competitive disadvantage or in conflict with regulatory and/or market expectations. However, sacred assumptions that prove incorrect create systemic risk precisely because they are widely held. Auditors must balance pragmatism with professional scepticism.

Key Expected Management Controls/ Possible Audit Tests

Governance review assesses whether the assumption is subject to any independent challenge. If documentation indicates the assumption is "non-negotiable" or "required by regulation," the auditor should verify whether the regulation actually mandates the specific assumption or merely permits it. Evidence review examines the evidence base - is it empirical or normative? Has it been updated recently? Are there emerging indicators that the assumption may no longer hold? Scenario analysis develops scenarios in which the sacred assumption fails and quantifies the impact. Disclosure review verifies that reliance on a sacred assumption is disclosed, particularly if the assumption is known to have limitations.

Gone Assumptions

Definition: A gone assumption is one that was once obvious or sacred but has become outdated or obsolete. The model stays. The world moves. Despite changes in market structure, regulatory environment, technology, or economic conditions, gone assumptions persist due to inertia, cost of change, or uncertainty about replacement assumptions.

Examples

Prior to the Global Financial Crisis, fair value measurements for certain illiquid structured products assumed continuous liquidity and price discovery. Post-crisis, this assumption is gone: liquidity is understood to be conditional and can evaporate under stress. However, some valuation models may still embed assumptions of continuous liquidity because updating the model is complex and costly.

Operational risk models developed before the digitisation of banking may assume that operational losses arise primarily from physical events (e.g. branch robberies, cheque fraud). This assumption is gone in an environment where cyber risk and digital fraud dominate operational loss profiles, yet legacy models may persist until replaced.

Cognitive Biases

Gone assumptions are primarily sustained by status quo bias, which is the preference for maintaining current practices over change, and the sunk cost fallacy: the belief that because resources were invested in building models based on the old assumption, those models should continue to be used. Recency bias may delay recognition that an assumption is gone if recent data still appears to support it, masking longer-term structural change.

Why This Status Matters for Audit

Gone assumptions introduce systematic risk because they misalign models with reality. Unlike speculative assumptions, which are known to be uncertain, gone assumptions may be embedded in models without recognition that they are no longer valid. Auditors must identify assumptions that reflect outdated environments and assess whether they remain fit for purpose.

Key Expected Management Controls/ Possible Audit Tests

Model inventory review identifies models developed or last substantially updated prior to known structural changes, assessing whether assumptions have been updated. Backtesting and outcome analysis compare model outputs to actual outcomes over recent periods - systematic bias or widening prediction errors may indicate gone assumptions. Environmental scanning reviews whether external conditions have changed materially since the model was developed. Replacement planning assesses whether management has a plan to update or replace identified gone assumptions; inaction may indicate governance weakness.

Speculative Assumptions

Definition: A speculative assumption is an acknowledged guess, adopted where insufficient evidence exists to support more confident assumptions. Speculative assumptions are necessary for forward-looking models - climate risk models, cyber risk models, or pandemic scenarios - where historical data is sparse or irrelevant.

Examples

In climate or cosmological models, speculative assumptions often dominate. While these disciplines often support their assumptions by proven laws of physics and their extrapolation, in a finance or business context, where agent behaviour is uncertain, speculative assumptions pose a major challenge.

In developing an expected credit loss estimate for a portfolio of loans to sectors exposed to climate transition risk, a bank may assume that carbon prices will increase by a specific percentage over the next decade. This assumption is speculative; carbon pricing regimes are uncertain and subject to political change. However, some assumptions are required to produce an estimate.

A strategic plan may assume that adoption of artificial intelligence will increase productivity by 1% per annum. This is speculative. AI's impact on productivity is uncertain and will vary by firm, sector, and implementation quality, but strategic planning requires assumptions about technology trends.

Cognitive Biases

Speculative assumptions are vulnerable to overconfidence bias (the tendency to overestimate the accuracy of one's predictions), optimism or pessimism bias (depending on the modeller's disposition and incentives), and anchoring on superficially similar historical precedents that may not be applicable.

Why This Status Matters for Audit

Audit cannot validate speculation. Its task is to ensure that it is clearly labelled as speculative, that sensitivity to the assumption is understood, that alternative scenarios have been considered, and that senior oversight is in place. Speculative assumptions must be regularly reviewed and updated as new data emerges. Consideration should be given to the amount and diversity of reviewers: while speculative assumptions require an expert making an informed prediction and should not be diluted by a myriad of different opinions, theoretical monocultures increase the risk of blind spots.

Key Expected Management Controls/ Possible Audit Tests

Labelling and disclosure review verifies that speculative assumptions are explicitly identified as such in model documentation and, where relevant, in financial statement disclosures or risk reporting. Scenario and range testing assesses whether management has developed alternative scenarios based on different speculative assumptions and whether the range of plausible outcomes is communicated. External anchoring review verifies that assumptions are benchmarked against external forecasts or expert analyses. Oversight review verifies that speculative assumptions are approved at an appropriately senior level, with committee minutes reflecting discussion of key speculative assumptions and their implications.

Summary table - Assumption Classification by Status

Status	Definition	Example	Key Cognitive Biases	Risk Level
Simplifying	Deliberately adopted to make a complex problem tractable by reducing dimensionality, aggregating heterogeneous elements, or abstracting away details believed not to materially affect the output for the intended purpose.	A bank assumes all borrowers with a given credit score exhibit the same probability of default when modelling expected credit losses on a credit card portfolio of millions of accounts. A trading desk assumes exchange rate returns are normally distributed for its VaR model, despite knowing the actual distribution has fatter tails.	Availability heuristic (gravitating toward familiar distributions); Anchoring bias (retaining original assumptions without re-examination in subsequent model iterations)	Low-Medium: Ubiquitous and necessary; risk arises when simplification introduces material distortion for the model's intended use (e.g., normal distribution adequate for 95% VaR but grossly inadequate for 97.5% Expected Shortfall).
Stating the Obvious	Reflects a widely accepted empirical regularity or market principle. Treated as self-evident. While it can be wrong under certain conditions, departing from it introduces new risks.	An impairment test assumes risk and return are positively correlated: higher-risk instruments require higher returns. A credit pricing model assumes obligors with higher credit ratings have lower probabilities of default, reflecting centuries of credit market history.	Conformity bias (social/professional pressure to align with consensus); Confirmation bias (preferentially noticing evidence supporting the obvious relationship and discounting anomalies)	Low under normal conditions; can become High during regime shifts. Dangerous when boundary conditions under which the assumption might fail are not identified or monitored.
Sacred	Normatively protected or accepted as absolute truth within a professional community or market context. Challenging it is perceived as naive or radical. Persists not because of evidence but because of institutional inertia and collective belief.	Pre-GFC: AAA-rated sovereign debt from developed economies was assumed risk-free for collateral and capital purposes (0% risk weight under Standardised Approach). Pre-GFC: residential mortgage defaults across US geographic regions were assumed to be largely uncorrelated, underpinning the entire securitisation market.	Conformity bias and authority bias (endorsed by regulators, rating agencies, academics); Groupthink (dissenting views suppressed); Fundamental attribution error (when sacred assumptions fail, failure attributed to external shocks rather than flawed assumptions)	Very High: Most difficult to challenge; creates systemic risk precisely because it is widely held. Incorrect sacred assumptions create simultaneous failures across institutions.

Status	Definition	Example	Key Cognitive Biases	Risk Level
Gone	Once obvious or sacred but has become outdated or obsolete due to changes in market structure, regulatory environment, technology, or economic conditions. Persists due to inertia, cost of change, or uncertainty about replacement.	Pre-GFC: fair value measurements for illiquid structured products assumed continuous liquidity and price discovery; post-crisis, liquidity is understood to be conditional. Operational risk models developed before digitisation assumed losses arise primarily from physical events (branch robberies, cheque fraud), now obsolete given cyber risk dominance.	Status quo bias (preference for maintaining current practices); Sunk cost fallacy (resources invested in old models justify continued use); Recency bias (recent data may mask longer-term structural change)	High: Introduces systematic risk because models are misaligned with reality. Unlike speculative assumptions, gone assumptions may be embedded without recognition they are no longer valid.
Speculative	An acknowledged guess, adopted where insufficient evidence exists to support more confident assumptions. Necessary for forward-looking models where historical data is sparse or irrelevant (climate risk, cyber risk, pandemic scenarios).	A bank assumes carbon prices will increase by a specific percentage over the next decade for climate transition risk ECL estimates. A strategic plan assumes AI will increase productivity by 1% per annum, despite uncertain and variable impacts.	Overconfidence bias (overestimating prediction accuracy); Optimism/pessimism bias (depending on modeller's disposition and incentives); Anchoring on superficially similar historical precedents	Medium-High: Cannot be validated; risk managed through labelling, sensitivity analysis, scenario alternatives, senior oversight, and regular review. Consideration needed on diversity of reviewers vs. theoretical monocultures.

Classification by Location

The location of an assumption describes where it sits within the modelling or decision-making process. Location determines how errors propagate through the model, whether they are foundational or introduced during a model run, and influences which audit techniques and types are most effective. Some location assumptions can only be changed with a new model; others can be changed in every run. The assumptions are listed in declining order of rigidity. Structural assumptions are often subject to testing in a model audit, whereas behavioural assumptions can be tested in standalone model audits, model-integrated audits, or business/financial audits.

Structural Assumptions

Definition: A structural assumption is built into the mathematical or statistical architecture of a model. It defines the functional form, distributional properties, and the model itself. Structural assumptions are often implicit, embedded in the choice of modelling technique.

Examples

A Gaussian copula model used for pricing CDOs assumes that default correlations remain constant across different economic states. This assumption is structural, inherent in the copula approach, and was central to pre-crisis CDO pricing. It failed catastrophically when correlations increased during the subprime mortgage crisis, fuelling the 2007/2008 financial crisis.

A VaR model that uses historical data and weighs recent periods more heavily than earlier periods makes two implicit assumptions: the past is representative of the future, and the recent past is more representative than the more distant past. The model may underestimate conditions not historically observed and inadequately account for medium-term economic cycles.

Error Propagation and Cognitive Biases

Structural assumptions affect all outputs of the model. If the structure is inappropriately specified, no amount of parameter calibration can meaningfully correct the error. Assuming a normal distribution for returns when the true distribution has fat tails will systematically underestimate extreme losses regardless of how accurately the mean and variance are estimated. Structural assumptions are also often shaped by availability heuristics. Modellers use techniques they are familiar with or that are standard in software packages. Confirmation bias sustains them, as subsequent validation focuses on parameters within the structure rather than questioning the structure itself.

Key Expected Management Controls/ Possible Audit Tests

Model architecture review traces the structural assumption through the model's logic, documenting functional forms, distributional assumptions, and causal relationships. Sensitivity testing across structures compares outputs from models with different structural assumptions (e.g. normal distribution vs. t-distribution for VaR). Historical validation backtests the model over periods that include stress conditions. Independent model review engages specialists to assess whether the chosen structure is appropriate for the application.

Contextual Assumptions

Definition: A contextual assumption defines the environment in which the model operates. It specifies external conditions such as market structure, legal and regulatory frameworks, macroeconomic state, or operational environment. Contextual assumptions are about the setting, not the model's internal logic.

Examples

A fair value model for sovereign bonds may assume that secondary markets remain open and liquid, allowing for price discovery. If markets close or become illiquid, the assumption fails, and fair value becomes difficult to determine.

A credit risk model may assume that central bank policy rates remain stable, implying no significant monetary policy intervention during the model horizon. If the central bank changes rates sharply - as occurred during the 2022–2023 tightening cycle - model outputs may diverge substantially from reality.

Error Propagation and Cognitive Biases

When contextual assumptions fail, models may not function at all, or may produce outputs entirely disconnected from reality. A credit risk model calibrated assuming economic growth may radically underestimate risk during a recession. These assumptions are shaped by normalcy bias, the expectation that the future will resemble the recent past, and recency bias, which causes modellers to overweight recent stable conditions and underweight historical periods of disruption.

Scenario Assumptions as a Special Case

A special type of contextual assumption is the scenario assumption, often used in stress tests. Scenario assumptions specify how the model - or the model output - will react to external shocks. Common scenario assumptions include economic downturns, rapid interest rate changes, or radical policy shifts such as the introduction of capital controls in a major economy. Because stress testing is central to prudential supervision, scenario assumptions warrant dedicated audit attention, particularly regarding the plausibility, internal consistency, and severity calibration of the scenarios used.

Key Expected Management Controls/ Possible Audit Tests

External alignment review compares contextual assumptions against current external data, regulatory changes, and macroeconomic forecasts. Trigger design review verifies that management has defined indicators signalling a change in context requiring model recalibration. Scenario analysis develops scenarios in which contextual assumptions fail and assesses whether the institution has contingency plans or alternative models. Governance review verifies that contextual assumptions are subject to periodic review and that responsibility for monitoring external conditions is clearly assigned.

Behavioural Assumptions

Definition: A behavioural assumption specifies how agents - customers, counterparties, market participants, employees - will act in response to conditions or stimuli. Behavioural assumptions are inherently uncertain because human behaviour is adaptive and context-dependent.

Examples

An expected credit loss model may assume that borrowers who become delinquent will cure (return to current status) at a certain rate based on historical data. If economic conditions deteriorate or forbearance programmes are introduced, borrower behaviour may change, invalidating the assumption.

A derivative pricing model may assume that higher-rated counterparties are more likely to exercise early termination options due to lower replacement costs, thus reducing effective maturity. If counterparty behaviour changes - due to changes in capital requirements or accounting treatment - the assumption may no longer hold.

Error Propagation and Cognitive Biases

Behavioural assumptions introduce uncertainty because behaviour can change due to incentives, information, regulation, stress, or human irrationality. Models assuming stable behaviour may fail when behavioural regimes shift. Feedback loops complicate matters: model outputs may influence agent behaviour, which in turn affects model accuracy. For example, published stress test results may alter bank behaviour. Projection bias (assuming others behave as we would) and status quo bias (assuming behaviours will remain stable) are also common.

The Double-Edged Sword of Herding

Herding bias may affect assumptions about market participants, though herding effects also occur in actual behaviour - assuming completely independent actions can be equally inaccurate. Bank runs like Northern Rock faced in 2007 are not uncommon and likely increasing with the immediacy of information propagation and widespread availability of online and mobile banking, as Silicon Valley Bank's (SVB) collapse in 2023 demonstrated. Assumptions about the likelihood, magnitude and velocity of herding should be considered whenever herding is identified as a risk factor.

Key Expected Management Controls/ Possible Audit Tests

Behavioural backtesting compares assumed behaviours against actual observed behaviours over time. Stress period analysis examines behaviour during past stress periods, as behaviour under stress often differs materially from normal conditions.

Feedback loop analysis assesses whether model outputs or management actions based on those outputs might influence agent behaviour. Segmentation review verifies that behavioural assumptions are appropriately segmented, as assuming all customers behave identically is often a simplifying assumption that may obscure material differences across segments.

Cross-Location Flexibility

One key feature of assumption location is that the same assumption can sit at different locations depending on how the model is built. For example, the simplifying behavioural assumption "all consumers in a country have identical preferences" corresponds to the common structural assumption "returns follow a normal distribution and relationships are linear." The corresponding contextual assumption could be defined as "the population is culturally and economically homogeneous, facing similar incomes and institutions."

Summary table - Assumption Classification by Location

Location	Definition	Example	Key Cognitive Biases	Risk Level
Structural	Built into the mathematical or statistical architecture of a model. Defines the functional form, distributional properties, and the model itself. Often implicit, embedded in the choice of modelling technique. Can only be changed with a new model.	The interest rate risk model uses a one-factor short-rate model (e.g., Hull-White) that assumes the entire yield curve can be described by movements in a single short-term rate. This structural choice determines all possible yield curve shapes the model can generate and excludes independent movements at different maturities.	Affects *all* outputs; no parameter calibration can correct structural misspecification. Shaped by availability heuristics (familiar techniques) and confirmation bias (validation focuses on parameters within structure rather than questioning structure itself).	Very High: Structural errors propagate through every output. If the structure is inappropriate, the model is fundamentally flawed regardless of calibration quality.
Contextual	Defines the environment in which the model operates. Specifies external conditions such as market structure, legal/regulatory frameworks, macroeconomic state, or operational environment. About the setting, not the model's internal logic.	The same interest rate risk model assumes central bank policy rates remain within a 0-5% corridor during the projection horizon, that secondary markets for government bonds remain open and liquid, and that no capital controls are imposed. These define the environment but not the model's mathematics.	When contextual assumptions fail, models may not function at all or produce entirely disconnected outputs. Shaped by normalcy bias (future resembles recent past) and recency bias (overweighting recent stable conditions, underweighting historical disruption).	High: Contextual failure can render entire models useless. Special attention needed for scenario assumptions used in stress tests (plausibility, internal consistency, severity calibration).

Location	Definition	Example	Key Cognitive Biases	Risk Level
Behavioural	Specifies how agents (customers, counterparties, market participants, employees) will act in response to conditions or stimuli. Inherently uncertain because human behaviour is adaptive and context-dependent. Can be changed in every model run.	The same interest rate risk model assumes that when rates rise above 3%, 15% of variable-rate mortgage holders will refinance to fixed rates within 6 months, and that depositors will not withdraw funds until the rate differential with competitors exceeds 100bps. These specify agent responses to the rate environment.	Behaviour can change due to incentives, information, regulation, stress, or irrationality. Feedback loops complicate matters (model outputs may influence behaviour). Shaped by projection bias (assuming others behave as we would) and status quo bias (assuming stable behaviour).	Medium-High: Inherently uncertain; risk increases with regime shifts. Herding effects (e.g., SVB, Northern Rock) demonstrate that behavioural assumptions about independent actions can fail catastrophically. Digital banking accelerates behavioural shifts.

Classification by Function

The function of an assumption describes its role in the model's logical flow - what it is doing within the reasoning process. Function determines the nature of assumption failure and appropriate testing approaches.

Baseline Assumptions

Definition: A baseline assumption establishes the foundational or reference state from which scenarios, forecasts, or valuations are developed. It defines the "starting point" or a condition "already there" before any analysis begins.

Examples

An impairment model may assume that the baseline economic scenario is consensus economist forecasts for GDP growth, unemployment, and interest rates. All probability-weighted scenarios are built relative to this baseline. The baseline assumption anchors the entire analysis.

A strategic planning exercise may assume continuation of the current business model and market structure as its baseline. If this baseline is overly optimistic (e.g. assuming current market share is sustainable when competition is intensifying), all subsequent analysis is skewed.

Failure Modes and Cognitive Biases

If the baseline is systematically biased, all scenarios will underestimate or overestimate risk. Baseline assumptions are particularly dangerous because they often receive less scrutiny than scenario assumptions - the baseline is perceived as "neutral" when it may embed significant judgement. These assumptions are rarely neutral. Optimism or pessimism bias, and normalcy bias can lead to baselines that extrapolate recent conditions rather than reflecting structural change.

Key Expected Management Controls/ Possible Audit Tests

Scenario comparison review compares the baseline against alternative baselines and external forecasts. Anchor risk review assesses how sensitive downstream metrics are to the baseline. External benchmarking compares against central bank projections, intergovernmental forecasts, or consensus surveys. Historical calibration assesses whether past baselines have proven reasonable; if they have historically been biased, governance should address this pattern.

Causal Assumptions

Definition: A causal assumption describes a directional relationship between variables: X causes Y, or X influences Y in a particular direction. Causal assumptions drive scenario design, stress testing, and predictive models.

Examples

An expected credit loss model may assume that rising unemployment causes increased default rates. If the causal link is weak or operates with long lags, the model may misjudge risk.

A pricing model may assume that increases in the policy interest rate cause increases in corporate bond yields. This assumption is grounded in monetary economics but can fail if other factors - quantitative easing, safe-haven demand - dominate.

Failure Modes and Cognitive Biases

Causal assumptions fail when the assumed relationship does not hold, reverses, or operates through unrecognised mechanisms. Correlation does not imply causation. Two variables may be driven by a common third factor, or the relationship may be coincidental. AI will happily mistake correlation for causation, an inevitable consequence of how it is built. These assumptions are vulnerable to illusory correlation, confirmation bias, and the post hoc ergo propter hoc fallacy.

Key Expected Management Controls/ Possible Audit Tests

Evidence review verifies that the causal link is, for example, supported by empirical analysis, economic theory, or subject matter expert judgement. Directionality testing examines historical data to assess stability and whether conditions exist under which the relationship reverses. Lag analysis assesses whether the model accounts for delays in the causal relationship. Alternative driver review challenges whether alternative causal structures could explain observed relationships.

Translational Assumptions

Definition: A translational assumption maps quantities, metrics, or concepts between different domains or measurement spaces. It converts X into Y, where X and Y are different types of variables.

Examples

A fair value model may translate observed volatility in the underlying asset into a price for a derivative. The Black-Scholes model translates equity volatility into option prices via a formula. The translational assumption is that the volatility measure accurately captures the uncertainty relevant to the derivative's payoff.

A credit model may translate credit spreads observed in the CDS market into default probabilities for counterparty exposure calculations. The translational assumption is that

CDS spreads are a reliable proxy for default risk, when in reality they also embed liquidity premia, market sentiment, and supply-demand dynamics.

Failure Modes and Cognitive Biases

Translational assumptions fail when the mapping does not preserve economic meaning or introduces systematic bias. For example, translating credit ratings into default probabilities assumes ratings accurately measure credit risk, which failed during the Global Financial Crisis when ratings for structured products proved systematically optimistic. The representativeness heuristic, i.e. assuming that a metric in one domain accurately represents a concept in another, and overconfidence in existing translation formulae are common biases.

Key Expected Management Controls/ Possible Audit Tests

Mapping validation compares the translation with backtested outcomes. Consistency review verifies that the same translational logic is applied consistently across portfolios. Information loss assessment evaluates what information is lost or distorted in the translation. Reverse translation testing, where feasible, translates back from Y to X to assess whether the original is recovered.

Transformational Assumptions

Definition: A transformational assumption specifies how a state changes into a modified version of itself over time or in response to events. It describes dynamics: X becomes X_1, which becomes X_2. Transformational assumptions govern path-dependent processes.

Examples

An interest rate model may assume that a surprise increase in central bank policy rates causes a parallel shift in the entire yield curve (all maturities shift by the same amount). In reality, curves may steepen or flatten, but the parallel shift assumption simplifies scenario design.

A pension liability model may assume that inflation increases pension obligations proportionally across all cohorts. If different cohorts have different inflation sensitivities (e.g. due to benefit caps), the assumption introduces error.

Failure Modes and Cognitive Biases

Transformational assumptions fail when dynamics are misspecified - transformations may be non-linear, asymmetric, or regime-dependent. Compounding errors over time can cause model paths to diverge materially from reality. Recency bias, linearity bias, and extrapolation bias are common influences.

Key Expected Management Controls/ Possible Audit Tests

Dynamic path testing runs scenarios over extended horizons to ensure transformations produce realistic time paths. Regime robustness review tests whether the transformation behaves reasonably across different regimes. Non-linearity testing assesses whether second-order effects matter. Compounding analysis traces how small errors accumulate over multiple periods.

Terminal Assumptions

Definition: A terminal assumption declares an end condition, boundary, or absorbing state where a process stops. It defines a technical boundary condition or a state beyond which the model no longer operates.

Examples

A VaR model calibrated to a one-year time horizon embeds the terminal assumption that losses beyond this horizon are outside the model's scope. If the institution's actual risk exposure extends beyond one year, as is typical for illiquid positions, the terminal assumption may cause material risk to go unmeasured.

A bank resolution plan may define capital thresholds below which the bank enters resolution and normal operations cease. The assumption is that the trigger is observable, that authorities will act, and that resolution mechanisms will function as designed.

Failure Modes and Cognitive Biases

Terminal assumptions can fail when boundary conditions are reached but the assumed end-state does not materialise as expected, or when triggers are poorly defined and terminal states are ambiguous. Operational executability is critical: a terminal assumption requiring rapid asset sales assumes buyers exist, which may not hold during stress or when instruments are illiquid. Optimism bias and normalcy bias are other common influences on terminal assumptions.

Key Expected Management Controls/ Possible Audit Tests

Boundary logic review assesses clarity and realism of triggers, thresholds, and end-states. Operational executability testing verifies that actions implied at the boundary are feasible given market conditions and operational capabilities. Historical precedent review examines whether terminal conditions have been reached previously and whether actual end-states resembled assumed ones. Contingency planning review verifies that management has tested contingency plans for reaching terminal states.

Summary table - Assumption Classification by Function

Function	Definition	Example	Key Cognitive Biases	Risk Level
Baseline	Establishes the foundational or reference state from which scenarios, forecasts, or valuations are developed. Defines the 'starting point' or a condition 'already there' before analysis begins.	An impairment model assumes the baseline economic scenario is consensus economist forecasts for GDP growth, unemployment, and interest rates, with all probability-weighted scenarios built relative to this baseline. A strategic plan assumes continuation of the current business model and market structure.	If the baseline is systematically biased, ALL scenarios underestimate or overestimate risk. Baselines receive less scrutiny than scenario assumptions despite embedding significant judgement. Heavily influenced by anchoring, optimism/pessimism bias, and normalcy bias.	High: Baseline errors cascade through all downstream analysis. Particularly dangerous because it is perceived as 'neutral' when it may embed significant judgement.
Causal	Describes a directional relationship between variables: X causes Y, or X influences Y in a particular direction. Drives scenario design, stress testing, and predictive models.	An ECL model assumes rising unemployment causes increased default rates. A pricing model assumes increases in policy interest rate cause increases in corporate bond yields (grounded in monetary economics but can fail when other factors like QE or safe-haven demand dominate).	Fails when assumed relationship does not hold, reverses, or operates through unrecognised mechanisms. Correlation does not imply causation. AI systems are particularly susceptible to drawing causal conclusions from correlations. Vulnerable to illusory correlation, confirmation bias, post hoc ergo propter hoc fallacy.	Medium-High: Causal misspecification leads to scenarios that do not reflect reality. Risk increases when causal chains are long or when relationships are regime-dependent.

Function	Definition	Example	Key Cognitive Biases	Risk Level
Translational	Maps quantities, metrics, or concepts between different domains or measurement spaces. Converts X into Y, where X and Y are different types of variables.	The Black-Scholes model translates equity volatility into option prices via a formula. A credit model translates CDS spreads into default probabilities for counterparty exposure calculations (CDS spreads also embed liquidity premia, market sentiment, and supply-demand dynamics).	Fails when mapping does not preserve economic meaning or introduces systematic bias. Credit ratings translated into default probabilities failed during the GFC when ratings proved systematically optimistic. Susceptible to representativeness heuristic and overconfidence in existing translation formulae.	Medium: Risk depends on fidelity of the mapping. Systematic bias in translation can propagate through all downstream calculations.
Transformational	Specifies how a state changes into a modified version of itself over time or in response to events. Describes dynamics: X becomes X1, which becomes X2. Governs path-dependent processes.	An interest rate model assumes a surprise increase in policy rates causes a parallel shift in the entire yield curve. A pension liability model assumes inflation increases obligations proportionally across all cohorts (ignoring different inflation sensitivities from benefit caps).	Fails when dynamics are misspecified: non-linear, asymmetric, or regime-dependent. Compounding errors over time cause model paths to diverge materially from reality. Susceptible to recency bias, linearity bias, and extrapolation bias.	Medium-High: Compounding errors make transformational assumptions increasingly dangerous over longer horizons. Small initial errors can produce large downstream deviations.
Terminal	Declares an end condition, boundary, or absorbing state where a process stops. Defines a technical boundary condition or a state beyond which the model no longer operates.	A VaR model calibrated to a one-year horizon embeds the assumption that losses beyond this horizon are outside scope (problematic for illiquid positions). A bank resolution plan defines capital thresholds below which the bank enters resolution, assuming triggers are observable and resolution mechanisms function as designed.	Fails when boundary conditions are reached but assumed end-state does not materialise or when triggers are poorly defined. Operational executability is critical: rapid asset sales assume buyers exist, which may not hold during stress. Susceptible to optimism bias and normalcy bias.	Medium: Risk is concentrated at boundaries. Low probability but potentially catastrophic impact when terminal conditions are reached and assumptions about end-states prove incorrect.

The Assumptions N-Space

Real-world assumptions are rarely reducible to a single dimension. An assumption in financial modelling typically has status (how entrenched it is), location (where it sits in the process), and function (what role it plays). Understanding all three dimensions in combination reveals risk that single-dimension analysis misses.

Why Combinations Matter

Consider the assumption that "housing prices will not decline simultaneously across all US geographic regions." This assumption was foundational to pre-crisis mortgage-backed securities pricing and rating.

- By status, this was a sacred assumption: embedded in decades of empirical data, supported by academic research, and reflected in rating agency methodologies.
- By location, it was contextual: an assumption about the macroeconomic and housing market structure, not about model internals.
- By function, it was causal: it posited that regional economic differences would cause housing markets to move independently, providing natural diversification.

The assumption's sacred status meant it received minimal challenge. Its contextual location meant that when the housing market environment changed, the assumption failed systemically. Its causal function meant that models based on independent regional price movements systematically underestimated tail risk.

How Audit Failures Occur

Audit failures in assumption review typically occur when one dimension is examined but others are ignored.

A status-only review verifies that assumptions are documented and approved but does not examine where they sit in the model architecture or what function they perform. Structural or causal assumptions embedded deep in model logic escape scrutiny. A location-only review maps assumptions to model components but does not assess entrenchment or function, allowing sacred assumptions to persist unchallenged because their status is not recognised. A function-only review tests causal relationships or sensitivity to baselines but does not assess whether assumptions are outdated or whether they propagate through structural architecture, missing gone assumptions or structural vulnerabilities.

Worked Examples

Risk-Free Rate (Historical)

Before the Global Financial Crisis, the assumption that AAA-rated sovereign debt from developed economies is risk-free was pervasive. By status, it was sacred, embedded in regulation, taught in universities, applied universally. By location, it was structural, built into capital models (zero risk weight), pricing models (risk-free rate as discount rate), and collateral frameworks. By function, it was a baseline, the risk-free rate served as an anchor for all relative pricing and risk measurement. Auditors examined whether the correct sovereign rates were used (parameter accuracy) but did not challenge the structural assumption that any rate could be "risk-free" or the sacred status that prevented meaningful challenge. When sovereign risk materialised during the Eurozone crisis, models failed across all three dimensions simultaneously.

Constant Default Correlation (Pre-Crisis CDOs)

Gaussian copula models assumed constant default correlations. By status, this was simplifying, modellers knew correlations varied but assumed constancy for tractability. By location, it was structural, embedded in the copula framework. By function, it was transformational, governing how shocks propagated through the portfolio. Some audits tested correlation parameter inputs but did not challenge the structural assumption that correlations remain constant. When correlations spiked in 2008, portfolios that appeared diversified based on normal-period correlations experienced simultaneous defaults.

Eternally Abundant Liquidity (Pre-Crisis)

The assumption that liquidity would always be available for high-quality assets underpinned funding strategies and asset valuations. By status, this was gone, it had been obvious for decades but became obsolete in 2008. By location, it was contextual, about market structure, not model internals. By function, it was terminal, liquidity defined boundaries, and when it disappeared, business models reliant on rollover funding reached terminal states. Audits focused on metrics (liquidity ratios) but did not challenge the contextual assumption that markets would remain open. The terminal function was not understood - liquidity was treated as a continuous variable rather than a binary state with catastrophic consequences at the boundary.

Toxic Combinations

Certain combinations of status, location, and function create disproportionate risk - not merely additive, but multiplicative. The worked examples above illustrate three patterns that recur across financial crises. Recognising these patterns enables auditors to prioritise where to focus scarce challenge capacity.

Sacred + Structural + Baseline is a common, but often the most dangerous combination. When an assumption is treated as unchallengeable (sacred), built into the architecture so it cannot be easily changed (structural), and serves as the anchor for all other analysis (baseline), the institution has a single point of failure that is often simultaneously invisible, inflexible, and fundamental. The risk-free rate assumption pre-crisis is the canonical example. Auditors encountering this combination should ensure these assumptions were reported to the most senior governance forum and require explicit scenario analysis of assumption failure, regardless of how "obvious" the assumption appears.

Gone + Contextual + Causal is also called a regime-change combination. The assumption once described reality but no longer does (gone), it defines the operating environment (contextual), and it specifies a cause-and-effect relationship that drives model behaviour (causal). When the world changes and the model might not, every causal chain in the model may be miswired. The pre-crisis liquidity assumption exhibited this pattern: the causal logic that "high-quality collateral ensures funding access" held for decades and then stopped holding. The audit response should be environmental scanning combined with explicit trigger design - defining observable indicators that would signal the assumption has become gone.

Speculative + Structural + Transformational is a combination susceptible to compounding errors. When a “guess” (speculative) is embedded in model architecture (structural) and governs how states evolve over time (transformational), small errors in the initial speculation compound through the model's dynamics, producing outputs that diverge progressively from reality. Climate risk models embedding speculative carbon price paths into structural economic relationships that transform asset values over decades can exemplify this pattern. The audit response requires dynamic path testing with wide uncertainty bands and explicit acknowledgement that model outputs at longer horizons are indicative rather than predictive.

Simplifying + Behavioural + Causal is the hidden-heterogeneity combination. When diverse agent behaviours are simplified into a single representative behaviour (simplifying + behavioural) and that behaviour results in specific outcomes (causal), the model may work on average but fail catastrophically when the simplified agents stop behaving uniformly. Retail deposit models assuming homogeneous depositor behaviour that "causes" stable funding exhibited exactly this pattern during SVB's collapse, when a concentrated depositor segment behaved very differently from the modelled average. The audit response is segmentation testing combined with concentration analysis.

Applying Multi-Dimensional Analysis

Effective assumption audit requires multi-dimensional analysis. This whitepaper uses three dimensions, hence the n-space can be described as a cube. However, additional dimensions may be helpful for some domains. Integrated Assessment Models (IAMs) used for climate forecasting, for example, combine economic, climate and technological elements. Therefore additional dimensions on the domain of the assumption can be beneficial. Conversely, very focused analyses on strategic decision-making in governance fora might not require location as a separate dimension. The approach, as well as the AI solution described in Part II, can be amended accordingly.

For each material assumption, the auditor should identify its type across all dimensions, assess the risk profile based on the combination, select audit methods appropriate to the matrix position, and document the matrix position in audit working papers. This enables consistent assessment across audits and supports identification of patterns, for example, multiple sacred assumptions in a single model may indicate groupthink.

Part II - Tools and Application

Core Tools

Sensitivity Analysis

Formal Definition: Sensitivity analysis quantifies the rate of change in a model's output with respect to changes in one (or a small set of) input assumptions, holding other assumptions constant. It answers the question: if assumption A changes by X%, how much does output Y change?

While the terms are often used interchangeably, sensitivity analysis and stress testing differ: stress testing often makes a holistic assessment of inputs and assumptions changing based on a scenario, while sensitivity analysis is a blunter instrument in which one or a few assumptions are moved up and down, sometimes in unrealistic magnitude, and most assumptions are deliberately held constant.

When Appropriate

Sensitivity analysis should be applied by the model owner or model risk management to all quantitative assumptions deemed material. It is particularly valuable for simplifying assumptions where the degree of simplification needs to be assessed, baseline assumptions that anchor scenarios, any assumption where uncertainty is recognised and quantifiable, and models used for decision-making where understanding assumption sensitivity informs risk appetite.

Limitations

Sensitivity analysis examines one assumption at a time (or small groups). It does not capture interaction effects or simultaneous changes in multiple assumptions. It assumes the model structure remains valid across the tested range. It does not address whether assumptions are fundamentally appropriate, only how sensitive outputs are to them.

As Audit Tool

Objective: While sensitivity analyses are primarily an expected control, they can be useful audit tools to assess whether model outputs are materially sensitive to key assumptions and whether management understands the impact of assumption uncertainty on decision-making.

1. Assumption Identification: Obtain model documentation and identify all quantitative assumptions. Prioritise based on materiality, uncertainty, and subjective judgement.
2. Range Definition: For each selected assumption, define a plausible range of variation referencing historical variability, expert judgement ranges, or scenario extremes.

3. Output Selection: Identify the model outputs that are decision-relevant (e.g. fair value, expected credit loss, VaR, capital requirement).
4. Execution: Vary each assumption across its defined range while holding others constant. Document the resulting output changes as both absolute and percentage change.
5. Tornado Chart Construction: Rank assumptions by their impact on output and present visually to illustrate which assumptions are most critical.
6. Multi-Way Sensitivity: For assumptions that are related or likely to move together (e.g. GDP growth and unemployment), perform multi-way sensitivity analysis to capture interaction effects.
7. Management Assessment: Interview model owners and decision-makers. Assess whether they are aware of the sensitivities, whether magnitudes align with expectations, and whether sensitivities inform their use of model outputs.

Typical Findings and Common Pitfalls

Common findings include management being unaware of which assumptions most affect outputs, sensitivity ranges not calibrated to plausible uncertainty, highly sensitive assumptions receiving inadequate governance, decisions made as if outputs are precise when sensitivity analysis reveals wide ranges, and sensitivity analysis existing but not being used to inform decision-making.

Common pitfalls include testing only parameter values while ignoring structural assumptions, using unrealistically narrow ranges that understate true sensitivity, performing sensitivity mechanically without connecting results to governance or decision processes, assuming linear sensitivity across the tested range when relationships may be non-linear, and failing to communicate sensitivity results to senior management.

Scenario Analysis

Formal Definition: Scenario analysis evaluates model behaviour and decision outcomes under a coherent set of simultaneously changed assumptions that describe a plausible future state of the world. Unlike sensitivity analysis, which isolates individual assumptions, scenario analysis asks: if the world looks like this, what happens to our model outputs and decisions?

Scenario analysis bridges sensitivity analysis (which tests individual assumptions in isolation) and backtesting (which tests against historical outcomes). It tests models against coherent, internally consistent narratives about future states - including states that have not occurred historically. This makes it the primary tool for testing contextual assumptions, sacred assumptions, and speculative assumptions, where historical data may be insufficient or misleading.

While in practice most scenario analyses are de facto stress tests, scenario analyses can also simulate different - not necessarily adverse - scenarios, and focus more on qualitative than quantitative outcomes.

Reverse Stress Testing aims to identify scenarios that would cause the institution to fail and assesses which assumption failures would trigger those scenarios. Reverse stress tests are not an “extra severe” stress test, but a scenario identification tool. If a model owner runs the same reverse stress test with the same scenario again and again, audit should raise this as a potential finding, as it defeats the purpose of a reverse stress test.

When Appropriate

Scenario analysis is appropriate when testing models against regime changes or structural shifts not captured in historical data, when evaluating the interaction effects of multiple assumptions changing simultaneously, when assessing contextual assumptions that define the operating environment, when testing sacred assumptions by constructing "what if this fails" narratives, and when evaluating speculative assumptions by exploring the range of plausible futures.

Limitations

Scenario analysis requires judgement in scenario construction, which itself introduces assumption risk. Scenarios may be too benign (anchored to recent experience), internally inconsistent (combining assumptions that cannot coexist), or insufficiently imaginative (failing to consider plausible but unprecedented combinations). The number of possible scenarios is effectively infinite, so selection bias is unavoidable - the question is whether the chosen scenarios cover the risk space adequately.

As Audit Tool

Objective: Like sensitivity analyses, material models should be stress tested by the model owner, model risk management or a separate stress testing or scenario analysis unit. While Audit can run its own scenario analyses, the primary focus should be assessing the robustness of the scenario analysis performed.

1. Scenario Inventory: Obtain the full set of scenarios used across risk management, capital planning, ICAAP/ILAAP, and financial reporting. Identify which assumptions each scenario tests.
2. Internal Consistency Review: For each scenario, assess whether the simultaneously changed assumptions are internally consistent. An economic downturn scenario that assumes rising unemployment but stable consumer spending is internally contradictory.
3. Severity Calibration: Assess whether scenario severity is appropriate. Compare against historical stress episodes, regulatory guidance, and peer practice. A scenario labelled "severe" should be at least as adverse as relevant historical crises.
4. Coverage Assessment: Map scenarios against the assumption taxonomy. Identify which assumption types (by status, location, and function) are tested by at least one scenario. Flag gaps - particularly untested sacred assumptions or speculative assumptions.
5. Output Analysis: Review how scenario results are reported, interpreted, and used. Assess whether results inform risk appetite, capital allocation, or management

actions - or whether they are produced for compliance without operational consequence.

Typical Findings and Common Pitfalls

Common findings include scenarios that are too similar to each other (testing variations on the same theme rather than fundamentally different states), sacred assumptions that are never tested because "they can't fail," internally inconsistent scenarios that produce misleading results, and scenario analysis performed for regulatory compliance without informing actual risk management.

Common pitfalls include anchoring scenarios to the most recent crisis rather than exploring different failure modes, treating scenario analysis as a one-time exercise rather than an ongoing process that evolves with the risk landscape, using scenarios designed by the same team whose assumptions are being tested (quis custodiet ipsos custodes?), and failing to consider second-order effects and feedback loops within scenarios.

Backtesting

Formal Definition: Backtesting compares model predictions or outputs against actual observed outcomes over a historical period. It assesses whether the model would have produced accurate forecasts or appropriate risk measures had it been applied to past data.

When Appropriate

Backtesting is appropriate when sufficient historical data exists to make comparisons meaningful, the model is used for prediction or risk measurement, outcomes are observable and measurable, and the environment has been relatively stable (or segmentation by regime is possible).

Limitations

Backtesting has critical limitations. Models designed to capture tail risk (e.g. 99.9% VaR) cannot be meaningfully backtested with limited data - if a 99.9% VaR has been breached once in ten years, this could indicate model accuracy or simply randomness. If the environment has changed structurally, historical backtesting may not predict future performance. Survivorship bias excludes failed entities where models may have failed most severely. Overfitting risk means models calibrated to historical data may backtest well by construction but fail prospectively.

Important: assumptions do not need to be "right" in an absolute sense. They need to be fit for purpose. Backtesting assesses fitness by examining whether assumptions produced outputs that supported sound decision-making, not whether they perfectly predicted outcomes. A model that is "wrong" but conservative may be fit for purpose if it prevents excessive risk-taking. A model that is "accurate" on average but misses tail events may not be fit for purpose if tail risk management is the objective.

As Audit Tool

Objective: Backtesting is an expected control for models. If the model owner or model risk management does not backtest models despite sufficient data, this can constitute a potential audit finding. Yet, independent backtesting can be a powerful audit tool, also to avoid backtests by the first and second line being inadequate.

1. Period Selection: Define the backtesting period. Include both normal and stress periods if possible. The period should be long enough to include meaningful variation but recent enough to remain relevant.
2. Outcome Definition: Define the observable outcome against which the model will be tested - actual defaults for credit models, actual profit and loss for VaR models, realised transaction prices for fair value models.
3. Prediction Extraction: Extract the model's predictions for the backtesting period. If the model has been recalibrated over time, use the predictions actually generated at each point (out-of-sample backtesting), not predictions generated retrospectively with current parameters.
4. Comparison Metrics: Calculate appropriate metrics - mean error, mean absolute error, and root mean square error for point predictions; coverage rates, magnitude of breaches, and clustering of breaches for distributional predictions; precision, recall, and false positive/negative rates for classification models.
5. Regime Segmentation: Segment results by regime (normal vs. stress, high vs. low volatility). Assess whether the model performs differently across regimes.
6. Bias Analysis: Assess whether errors are random or systematic. Consistent overestimation or underestimation indicates bias.
7. Assumption Attribution: For material errors, trace back to the underlying assumptions. Which assumptions contributed to the error? Has management updated those assumptions or documented why they remain appropriate?
8. Documentation Review: Verify that backtesting results are documented, reviewed by appropriate governance fora, and used to inform model updates or usage restrictions.

Typical Findings and Common Pitfalls

Common findings include backtesting performed but results not used to inform model governance, models performing well in normal periods but failing during stress without usage restrictions, systematic bias existing but not corrected or explained, and sample sizes too small to draw meaningful conclusions with management treating results as definitive.

Common pitfalls include applying backtesting to models where outcomes are unobservable, expecting perfect accuracy and treating any deviation as failure, using backtesting to validate rare-event models where statistical power is insufficient, failing to account for regime changes, and backtesting using in-sample data that produces artificially favourable results.

Applications

This section applies the assumption taxonomy and audit tools to two domains that illustrate different challenges: derivative pricing and Level 3 instrument valuation, where the challenge is model complexity and market-based assumptions; and operational deposit modelling, where the challenge is behavioural uncertainty and recent regime changes. The pricing section provides a framework-level application; the deposit section provides a detailed audit programme reflecting post-SVB regulatory expectations.

Pricing: Derivatives and Level 3 Instrument Pricing[1]

Typical Pricing Assumptions

Derivative pricing and fair value measurement of Level 3 instruments involve numerous assumptions: market structure assumptions (liquidity is continuous, markets are complete, no arbitrage opportunities exist), volatility assumptions (historical volatility predicts future volatility, volatility is constant or mean-reverting), correlation assumptions (correlations between underlyings remain stable), credit assumptions (counterparty default probabilities, recovery rates, exposure at default), funding assumptions (institutions can borrow and lend at consistent rates), and behavioural assumptions (counterparties exercise options rationally, early termination probabilities are stable).

Assumption Matrix Position

By status, pricing assumptions range from simplifying (constant volatility) to sacred (no-arbitrage) to gone (perpetual liquidity). Credit Valuation Adjustment (CVA) and Debit Valuation Adjustment (DVA) are recent additions that acknowledge previously gone assumptions about counterparty risk. By location, structural assumptions include distributional forms; contextual assumptions include market accessibility and legal enforceability; behavioural assumptions include counterparty exercise behaviour. By function, baseline assumptions define reference prices or rates; causal assumptions link market factors to price changes; translational assumptions convert observed market data into model inputs; transformational assumptions govern price evolution over time.

How Audit Should Test Pricing Assumptions

1. Model Selection Justification: Verify that the choice of pricing model (e.g. Black-Scholes vs. binomial vs. Monte Carlo) is appropriate for the instrument's characteristics.

2. Input Assumption Validation: For each model input, verify source of observable inputs, verify estimation methodology for unobservable inputs, and assess consistency with observable proxies.

3. XVA Framework Review: Verify that appropriate valuation adjustments are applied. For CVA, assess whether counterparty default probabilities and exposures are reasonable. For

[1] This section is written in an accounting-standard-agnostic manner. Auditors are advised to verify current and upcoming changes to IFRS and local GAAP rules applicable to the entity’s financial reporting regime

DVA, verify calculation methodology. For FVA, assess whether funding cost assumptions reflect the institution's actual funding profile.

4. Consensus Price Risk Assessment: For Level 3 instruments priced using consensus data, identify the number of contributing institutions, assess whether internal pricing is consistent with consensus, and assess risk of contamination if few contributors use similar models.

5. Sensitivity and Scenario Analysis: Vary key pricing assumptions (volatility, correlations, credit spreads) to assess output sensitivity. Verify that management understands valuation uncertainty.

6. Independent Price Verification: Where feasible, obtain independent pricing from third parties or use alternative models. Investigate material differences.

7. Profit and Loss Attribution: For traded portfolios, verify that daily P&L is explained by market moves and known model sensitivities. Unexplained P&L may indicate model issues or inappropriate assumptions.

Typical Failure Patterns

Over-reliance on historical data during regime changes causes volatility assumptions based on recent benign periods to underestimate risk when markets become stressed. Correlation failure during stress eliminates diversification benefits as correlations converge toward 1 during crises. Circular consensus occurs when institutions submit internal prices to consensus services, then use consensus prices to validate their internal prices, creating a self-reinforcing loop that may embed systematic error. Counterparty credit assumptions that become sacred - such as the pre-crisis assumption that counterparty risk was negligible for short-dated or highly-rated counterparties - prove costly when they fail. DVA recognition creates situations where institutions report gains as their own creditworthiness deteriorates, which - while following accounting logic - may obscure true financial condition.

Operational Deposit Modelling: A Detailed Application[2]

The operational deposit modelling section provides a more granular application of the framework, reflecting the post-SVB regulatory focus on deposit stability assumptions and the particular challenges of modelling depositor behaviour in a digital banking environment. The level of detail here is deliberately greater than the pricing section, serving as an illustrative deep-dive that demonstrates how the taxonomy translates into a complete audit programme.

[2] This section is written in a regulation-agnostic manner. Auditors are advised to review current and upcoming regulatory requirements. In particular for EU entities, upcoming changes in EU’s COREP requirements should be considered.

Account Segmentation and Operational Base Assumptions

Segmentation models embed causal assumptions (account characteristics determine operational behaviour and stability), baseline assumptions (through-the-cycle or point-in-time operational balances as reference points), and behavioural assumptions (depositor response to rate changes, competitive offers, or confidence crises). By status, segmentation often contains obvious assumptions (high-turnover accounts are less stable) but may contain gone assumptions if calibrated to pre-SVB environments or if they fail to reflect digital banking realities.

Stability and Runoff Rate Assumptions

Stability assumptions span contextual assumptions (regulatory treatment under LCR, legal and operational withdrawal constraints, switching costs, deposit insurance thresholds, digital banking infrastructure), transformational assumptions (stable balances decline by certain percentages under stress scenarios), and structural assumptions (runoff distributions such as exponential decay or survival models that may not capture digital-era withdrawal speeds, social media-accelerated runs, or concentrated segment contagion). Trade cycle assumptions reflecting client-specific business model characteristics should align with EBA 2025 guidance.

Interest Rate Sensitivity and Concentration Risk

Behavioural assumptions about depositor migration to higher-yielding alternatives when rate differentials exceed behavioural thresholds (deposit rate elasticity) are central. Translational assumptions about current beta coefficients translating to stressed sensitivities via repricing lags (typically 3–9 months) are regime-dependent across low/high and rising/falling rate environments. Contextual assumptions about deposit concentrations (individual depositors, industries, uninsured balances) behaving independently under stress may understate correlated withdrawal risk.

Stress Testing and Scenario Design

Stress testing for deposit models involves scenario assumptions (projecting deposit flight paths, rate wars, confidence crises not observed in the current market structure), baseline assumptions (starting deposit mix and concentration levels from which stress scenarios diverge), and causal assumptions (how rate shocks, reputational events, or competitive actions propagate through the deposit base). The SVB lesson is that interest rate shocks, deposit withdrawals, and asset mark-to-market losses occur simultaneously - considering multi-dimensional stress is essential.

Common issues include recency bias in scenarios (stress scenarios resembling the last deposit crisis rather than exploring different failure modes), underestimation of feedback loops (first-order effects modelled but second-order effects such as social media acceleration and cross-segment contagion omitted), benign baselines that do not reflect early warnings, inadequate withdrawal speed assumptions (standard LCR assumes 5% stable retail deposit runoff over 30 days in many jurisdictions; SVB experienced over 20% in a single day), and optimistic stability factors (NSFR treatment of uninsured wholesale deposits in some jurisdictions as 50% for non-financial counterparties "available" may be unrealistic for concentrated, rate-sensitive, or VC-funded deposit bases).

Distinguishing Model Error, Assumption Risk, and Governance Failure

Audit must distinguish three failure modes because remediation differs. Model error comprises technical implementation failures - coding errors, incorrect formulae, segmentation mapping mistakes, data processing errors - identifiable through code review, reconciliation, and independent recalculation. Assumption risk means inappropriate, outdated, or inadequately justified assumptions causing misleading outputs even when implementation is technically correct, identified through the assumption taxonomy, behavioural backtesting, and regime-dependency testing. Governance failure means inadequate oversight, absent or ineffective challenge of assumptions, poor communication of limitations, or failure to update when conditions change materially, identified through documentation review, committee minute analysis, and management interviews.

Remediation for model error is technical: fix the code, correct formulae, repair data feeds. Remediation for assumption risk requires updating assumptions for current conditions, recalibrating for relevant regimes, or restricting model usage with explicit overlays. Remediation for governance failure requires strengthening oversight, improving documentation, enhancing challenge processes, and establishing trigger-based reviews.

Operational Deposit Modelling Audit Programme

1. Assumption Inventory: For each segment, create a comprehensive inventory of assumptions across status, location, and function dimensions - particularly operational definitions, stability thresholds, behavioural triggers, trade cycle justification, concentration treatments, and regime-dependent behaviours.

2. Regulatory Compliance and Trade Cycle Review: Verify operational deposit identification complies with EBA guidance on legal and operational constraints. Assess whether trade cycle length reflects client-specific business models and whether excess operational deposit calculations use prudent statistical measures.

3. Behavioural Assumption Testing: Compare assumed behaviours (runoff rates, rate sensitivities, seasonal patterns) against historical actuals segmented by rate environments, competitive conditions, and economic cycles. Backtesting must cover multiple regimes including digital-era withdrawal speeds.

4. Concentration and Segmentation Testing: Analyse deposit base diversification and identify material concentrations. Verify account classification into operational vs. non-operational is defensible and validated against observed stability. Test whether "operational base" accounts demonstrate operational behaviour through transaction pattern analysis.

5. Scenario Plausibility Assessment: Verify stress scenarios are internally consistent and severity aligns with regulatory expectations and historical precedents (SVB, Signature Bank). Test combined stresses, speed scenarios (1-day, 1-week, 30-day), segment-specific shocks, and feedback loops.

6. Model Limitation Documentation: Verify documentation states what the model does not capture (social media runs, cross-segment contagion, post-SVB withdrawal speeds, concentration risk adjustments), which risks it does not address, and under which conditions it may be unreliable. Documentation must state calibration period and whether it includes stress episodes.

7. Governance and Output Usage Review: Review committee minutes, validation reports, and audit findings to assess whether assumption challenges are substantive or perfunctory. Verify model outputs are used consistently with limitations: if the model underestimates digital-era flight speeds or calibration predates regime shifts, verify management applies buffers or overlays. Assess integration with concentration risk limits, capital planning, ALM, and liquidity planning.

AI-Assisted Assumption Identification and Classification

Most audit teams are still manually hunting through model documentation for assumptions, which is exactly as tedious as it sounds. There is a better approach: building an AI assistant[3] that does the heavy lifting of identifying and classifying assumptions systematically across the model inventory.

What It Does

The concept is straightforward. Feed the AI model documentation - validation reports, committee papers, methodology guides - and it does four things.

First, it finds the assumptions. It scans documents looking for statements that sound like assumptions: assertions about relationships, behaviours, or market conditions that are not directly verifiable, statements such as "counterparties behave rationally" or "correlations remain stable during stress."

Second, it classifies them. Using the taxonomy framework (status, location, function), it categorises each assumption through keyword matching, contextual analysis, and pattern recognition. Classification is the key step and requires training on a key set of assumptions. Three different approaches can be used to train an AI system on assumption classification:

- By dimension: Each assumption is separately assessed across each dimension, and then the result aggregated. This provides robustness of the output and a larger training set, as each dimension contains hundreds of examples.

[3] As described in Part I, AI models are susceptible to bias and model risk as any other models. Any reliance on AI and its output should undergo thorough review by a human.

- By n-space position: The AI system is trained on each position in the n-space. This enables targeted identification of toxic combinations, but relies on smaller sample sets per n-space position.
- By definition: The AI system is not trained on existing assumption, but on the definitions, which forces it to reason what the adequate classification is, but does not necessarily learn from real-life examples.

A freely available prototype[4] of a classification system by dimension can be found in Google Gems.[5] Other than the full AI, this freely available prototype does not scan texts to identify assumptions, but will classify them along the three dimensions and provide audit guidance.

Each dimension is trained on more than 400 assumptions from seven separate domains ranging from finance to climate science and medicine. The n-space position as well as the definition were built in as control layer. The combination of the separate assessment by dimension provides the n-space position, the AI provides an accuracy confidence score that is determined by the n-space position method and the assessment via definition. The training sets for each dimension, n-space position and definition based on this whitepaper's methodology are freely available upon request.

Third, it tracks them over time. It builds a registry showing which assumptions appear across multiple models, which ones have been unchanged since 2015, which ones keep changing. This is where patterns that matter start becoming visible.

Fourth, it scores the risk. Not all assumptions are equally risky. The system flags assumptions that are material, uncertain, poorly governed, or have a track record of validation issues. This is what tells you where to spend your time.

Beyond Speed: Completeness, Consistency, and Pattern Recognition

The practical benefit is not just speed - though speed is gained. The primary value is completeness and consistency. Manual reviews inevitably miss things, especially implicit assumptions buried in model structure. An automated system does not get tired or skip boring sections. More importantly, it applies the same classification logic everywhere, which means an enterprise-wide view of what assumptions the institution is actually sitting on.

This enterprise view enables a class of questions that manual reviews cannot practically answer. Are certain cognitive bias patterns showing up repeatedly across model families? Is governance consistently blind to specific assumption types - for example, are sacred assumptions systematically absent from challenge logs? Are we inheriting decades-old assumptions that nobody has challenged because they predate the current model validation framework? Are the same assumptions embedded in multiple models without

[4] The prototype serves as a demonstration and idea generation tool. It must not be considered as an approved third-party solution or external data provider. The use of any output of the tool for internal use must be subject to rigorous review and challenge by accountable individuals.

[5] Requires an active Google Account. Ensure to log on before clicking the link to avoid being routed to Gemini's front page.

anyone recognising the concentration risk? These are the questions that transform assumption audit from a compliance exercise into a genuine risk management tool.

Toxic Combination Detection

The AI system can be specifically trained to identify the toxic combinations described in the n-space section. By classifying each assumption across all three dimensions simultaneously, it can flag combinations that warrant escalation - sacred-structural-baseline assumptions that represent single points of failure, or gone-contextual-causal assumptions that indicate unrecognised regime changes. This pattern detection across the full model inventory is where the automated approach most clearly outperforms manual review.

Limitations and Human Oversight Requirements

Automated systems support but do not replace human judgement. Determining whether an assumption is fit for purpose requires understanding the model's use case, its organisational context, and the competitive and regulatory environment - capabilities that current AI systems do not possess. The boundaries between assumption types can be ambiguous, and human review is required for edge cases and for validating the system's classifications. Novel assumptions arising from new products, markets, or risks will not match existing patterns and require human identification.

There is also a meta-risk: the assumption classification system itself embeds assumptions about what constitutes an assumption, how categories should be defined, and which patterns indicate risk. An AI system that applies the taxonomy mechanically may develop blind spots of its own - for example, consistently classifying an assumption as "simplifying" when a human reviewer with domain expertise would recognise it as "gone." Regular calibration of the AI system's outputs against expert human judgement is essential, particularly when the external environment changes.

Building Internal Capabilities

Institutions need not rely on external vendors to develop assumption management tools. Similar capabilities can be built internally using natural language processing libraries, knowledge graphs to map relationships among assumptions, models, and governance artefacts, machine learning classifiers trained on institution-specific model documentation, and integration with existing model inventory and governance platforms.

The key challenge is not technological but organisational. Assumption management must be integrated into model risk management and audit workflows rather than treated as a standalone analytical exercise. This requires buy-in from model owners (who must accept that their assumptions will be systematically catalogued and challenged), model validation (who must incorporate the taxonomy into their review frameworks), and senior management (who must be willing to act on the enterprise-wide patterns the system reveals).

Conclusion

Assumptions have become a key audit challenge. They are upstream of models, meaning weak assumptions can invalidate even strong mathematics. They are difficult to challenge without a structured approach. The taxonomy presented in this whitepaper, classifying assumptions by status, location, and function, provides a framework to address this challenge. It enables auditors to systematically identify assumptions, assess their risk profile through multi-dimensional analysis, select appropriate testing methods, and evaluate governance adequacy.

The key principle is fitness for purpose. Assumptions must be appropriate for their intended use, not "right" in an absolute sense, unless they are demonstrably inaccurate and not useful, i.e., detrimental to the model's intended outcome. The auditor's responsibility is to ensure that management can articulate why their chosen assumptions are suitable, that alternative assumptions and sensitivities to changes have been considered, and that limitations are clearly communicated.

The Layered Beliefs Framework illustrates that financial reality is constructed from interconnected layers: assumptions, models, consensus, and markets. Audit assurance is placed on the robustness of the processes that generate, translate, aggregate, and apply these beliefs. The three tools covered in Part II of this whitepaper, sensitivity analysis, scenario analysis, and backtesting, provide practical instruments applicable across financial statement audits, model validation, and internal audit. Auditors should, however, consider additional tools as expected controls, or audit techniques as required.

The importance of assumption audits will only increase. Financial markets continue to evolve with greater model complexity. AI-driven decision-making and emerging risk categories that lack historical precedent further increase such complexity. The toxic combinations identified in this paper offer auditors a guidepost for identifying where assumption risk concentrates. AI-assisted assumption management can support enterprise-wide visibility. Human judgement, however, remains essential for determining fitness for purpose. AI-assisted solutions, while helpful, introduce additional risks and require adequate oversight.

Crucially, these principles extend beyond financial services to any domain reliant on expert judgement, predictive models, or forward-looking estimates. Examples include medicine and healthcare, climate risk assessments, infrastructure planning, or any other field where decisions rest on beliefs about unknown relations or an uncertain future. Audit's role is to ensure that uncertainty is acknowledged, assumptions are made transparent and are justified, and decision-makers fully understand the limitations of the information they rely on.

Appendices

Appendix I - Assumptions N-Space: All 75 Combinations

Status	Location	Function	Example	Audit Considerations	Key Audit Steps
Simplifying	Structural	Baseline	VaR model using normal distribution as the baseline statistical framework, simplifying fat-tailed return distributions for computational tractability.	Low-Medium. Simplification in model architecture anchoring all risk measurement. Test whether simplification materially affects baseline calibration.	Sensitivity analysis comparing simplified vs. complex baseline; benchmark against alternative distributions.
Simplifying	Structural	Causal	Linear regression model assuming all customer segments respond identically to price changes, simplifying heterogeneous price elasticities into a single coefficient.	Medium. Simplified causal structure may miss non-linear or segmented relationships.	Segmentation testing; compare linear vs. non-linear model outputs; assess residual patterns.
Simplifying	Structural	Translational	Credit model using a single mapping table to translate all external ratings into internal PD estimates, ignoring sector-specific default patterns.	Medium. Simplified translation embedded in architecture may introduce systematic bias for certain sectors.	Backtesting translation accuracy by sector; compare against sector-specific mappings.
Simplifying	Structural	Transformational	Interest rate model assuming parallel yield curve shifts for all stress scenarios, simplifying the complex dynamics of curve steepening and flattening.	Medium. Simplified dynamics may compound over time, particularly for long-dated instruments.	Compare parallel shift outputs vs. non-parallel scenarios; assess impact on long-dated positions.
Simplifying	Structural	Terminal	Recovery rate model applying a flat 40% recovery assumption to all defaulted bonds regardless of seniority, collateral, or jurisdiction.	Low-Medium. Simplified terminal condition; risk depends on portfolio diversity.	Segment recovery assumptions by seniority and collateral type; backtest against realised recoveries.
Simplifying	Contextual	Baseline	ECL model using a single macroeconomic forecast as the baseline scenario, simplifying the range of possible economic environments.	Medium. Oversimplified baseline context may bias all probability-weighted scenarios.	Compare against multiple external forecasts; test sensitivity of ECL to baseline variation.

Status	Location	Function	Example	Audit Considerations	Key Audit Steps
Simplifying	Contextual	Causal	Stress test assuming a single transmission channel from GDP decline to credit losses, simplifying the complex causal pathways through employment, asset prices, and confidence.	Medium. Simplified causal context may miss important secondary transmission channels.	Map all plausible transmission channels; assess which are captured vs. omitted; test omitted channels.
Simplifying	Contextual	Translational	FX model using a single exchange rate regime assumption to translate foreign exposures, ignoring the possibility of capital controls or currency pegs breaking.	Medium. Simplified contextual translation may fail during regime changes in currency markets.	Scenario test under alternative FX regimes; assess hedging effectiveness under different translations.
Simplifying	Contextual	Transformational	Liquidity model assuming a constant market impact function for asset sales during stress, simplifying the non-linear relationship between selling pressure and price impact.	Medium-High. Simplified transformation of market conditions during stress may severely underestimate liquidation costs.	Test non-linear market impact functions; compare with empirical evidence from past fire sales.
Simplifying	Contextual	Terminal	Business continuity model assuming a fixed maximum disruption period of 30 days for operational events, simplifying the range of possible disruption durations.	Low-Medium. Simplified terminal boundary for operational context; test whether the cap is realistic.	Review historical disruption durations; scenario test extended disruptions beyond the assumed cap.
Simplifying	Behavioural	Baseline	Customer attrition model using portfolio-average churn rates as the baseline, simplifying diverse customer loyalty behaviours into a single metric.	Medium. Simplified behavioural baseline may mask significant variation across customer segments.	Segment churn by customer type; compare segment-level vs. portfolio-level baselines.

Status	Location	Function	Example	Audit Considerations	Key Audit Steps
Simplifying	Behavioural	Causal	Retail deposit model assuming homogeneous depositor behaviour that causes stable funding, simplifying diverse depositor segments into a single representative agent with uniform response patterns.	TOXIC COMBINATION: Simplifying-Behavioural-Causal (Hidden-Heterogeneity Combination). Diverse agent behaviours simplified into single representative behaviour that causes specific outcomes. Works on average but fails catastrophically when simplified agents stop behaving uniformly (e.g., SVB concentrated depositor segment).	Segmentation testing combined with concentration analysis. Test whether sub-populations behave materially differently. Assess concentration risk in depositor/counterparty base.
Simplifying	Behavioural	Translational	Prepayment model translating interest rate changes into a single aggregate prepayment speed, simplifying the diverse refinancing behaviours across borrower segments.	Medium. Simplified behavioural translation may underestimate prepayment risk for specific cohorts.	Backtest prepayment speeds by borrower segment; compare aggregate vs. segmented translation accuracy.
Simplifying	Behavioural	Transformational	Deposit model assuming all depositors shift balances at the same constant rate when rates change, simplifying the heterogeneous speed and magnitude of depositor responses.	Medium-High. Simplified behavioural dynamics may understate the speed of deposit flight in digital banking environments.	Segment depositor behaviour by type; test accelerated withdrawal scenarios; compare with SVB-era data.
Simplifying	Behavioural	Terminal	Loan workout model assuming all defaulted borrowers either cure or are written off within 12 months, simplifying the range of resolution timelines.	Low-Medium. Simplified terminal state for behavioural outcomes; test whether boundary is realistic across segments.	Backtest actual resolution timelines by segment; assess whether the 12-month cap holds under stress.

Status	Location	Function	Example	Audit Considerations	Key Audit Steps
Obvious	Structural	Baseline	Credit scoring model using credit bureau scores as the foundational risk ranking mechanism, based on the obvious empirical regularity that past payment behaviour predicts future behaviour.	Low-Medium. Well-established baseline but may fail for populations with thin credit files or during structural economic shifts.	Boundary analysis: identify conditions where the score-default relationship weakens; backtest across economic cycles.
Obvious	Structural	Causal	Asset pricing model built on the assumption that higher risk requires higher expected return, structurally embedding the risk-return trade-off into the pricing framework.	Low under normal conditions; elevated during flight-to-quality episodes where risk and return decouple.	Historical backtesting during stress; opposite assumption testing; monitor for anomalous periods.
Obvious	Structural	Translational	Bond pricing model translating credit ratings into yield spreads using historically observed rating-spread mappings, based on the obvious relationship between creditworthiness and borrowing cost.	Low-Medium. Well-established translation; risk arises during rating transition periods or when spreads decouple from fundamentals.	Backtest rating-spread mappings across market cycles; assess spread decomposition beyond credit risk.
Obvious	Structural	Transformational	Economic capital model assuming that diversification reduces portfolio risk over time as the number of independent exposures increases, based on the law of large numbers.	Low under normal conditions; High when correlations spike and diversification benefits evaporate.	Stress-test correlation assumptions; compare diversified vs. concentrated portfolio behaviour in crises.
Obvious	Structural	Terminal	Insurance claims model with a policy limit serving as the maximum payout boundary, based on the obvious contractual cap on losses.	Low. Contractually defined terminal condition; risk only if policy wordings are ambiguous or contested.	Review policy language for ambiguity; assess litigation risk around coverage disputes.

Status	Location	Function	Example	Audit Considerations	Key Audit Steps
Obvious	Contextual	Baseline	Sovereign risk model using developed-market government bond yields as the risk-free baseline, based on the widely accepted principle of sovereign creditworthiness hierarchy.	Low-Medium. Standard practice; risk elevated during sovereign debt crises when 'risk-free' assets become risky.	Monitor sovereign CDS spreads; scenario test sovereign stress; review for Eurozone-crisis-type precedents.
Obvious	Contextual	Causal	Monetary policy transmission model assuming that central bank rate increases cause mortgage rates to rise, based on well-established monetary economics.	Low-Medium. Generally holds but transmission mechanisms can vary by market structure and competitive dynamics.	Assess pass-through rates historically; test for lag variability; identify conditions where transmission weakens.
Obvious	Contextual	Translational	IFRS 13 fair value model translating observable market prices into fair values, based on the obvious principle that traded prices in active markets represent fair value.	Low for liquid instruments; Medium-High for illiquid instruments where price discovery is limited.	Assess market liquidity; review bid-ask spreads as indicator of price reliability; compare across pricing sources.
Obvious	Contextual	Transformational	Inflation model assuming that monetary expansion eventually leads to price increases, based on the widely accepted quantity theory of money relationship.	Medium. Long-run relationships generally hold but timing and magnitude are highly uncertain (as demonstrated post-2008 QE).	Test lag structures; compare monetary expansion periods with inflation outcomes; assess structural breaks.
Obvious	Contextual	Terminal	Deposit insurance model assuming government guarantees will be honoured up to the statutory coverage limit, based on the obvious expectation of sovereign backstop.	Low under normal conditions; risk rises during systemic crises when guarantee credibility is tested.	Scenario test sovereign fiscal capacity; review historical instances of deposit guarantee activation.
Obvious	Behavioural	Baseline	Retail credit model using historical average default rates as the baseline for customer repayment behaviour, based on the obvious pattern that most borrowers repay their debts.	Low-Medium. Standard baseline; risk when economic conditions diverge materially from historical norms.	Backtest baseline against multiple economic cycles; assess whether historical average remains representative.

Status	Location	Function	Example	Audit Considerations	Key Audit Steps
Obvious	Behavioural	Causal	Collections model assuming that increased collection efforts cause higher recovery rates, based on the obvious relationship between effort and outcome in debt recovery.	Low-Medium. Generally holds but effectiveness varies by borrower segment, economic conditions, and regulatory constraints.	Backtest collection effectiveness by segment; assess diminishing returns; test regulatory impact on collection strategies.
Obvious	Behavioural	Translational	Customer lifetime value model translating purchase frequency into retention probability, based on the obvious principle that frequent buyers are more loyal.	Low-Medium. Generally valid translation; risk for subscription models where frequency may not reflect loyalty.	Validate frequency-loyalty relationship by customer segment; assess for model-specific exceptions.
Obvious	Behavioural	Transformational	Credit migration model assuming that borrowers gradually deteriorate through rating grades before default, based on the obvious pattern of progressive financial distress.	Medium. Generally holds but sudden defaults (fraud, operational failures) bypass the gradual deterioration path.	Backtest migration patterns; assess frequency of sudden defaults vs. gradual deterioration; test for jump-to-default risk.
Obvious	Behavioural	Terminal	Loan model assuming borrowers in default either cure, restructure, or are written off within a standard workout period, based on typical resolution patterns.	Low-Medium. Standard terminal framework; risk during systemic crises when workout timelines extend significantly.	Backtest resolution timelines across economic cycles; stress test for extended workout periods.
Sacred	Structural	Baseline	Pre-GFC assumption that the risk-free rate derived from AAA sovereign debt was immutable, built into every pricing and capital model as the foundational discount rate anchor.	TOXIC COMBINATION: Sacred-Structural-Baseline. Single point of failure that is simultaneously invisible, inflexible, and fundamental. When an unchallengeable assumption is built into architecture and serves as anchor for all analysis, failure is catastrophic and systemic. The risk-free rate assumption pre-crisis is the canonical example.	Escalate to the most senior governance forum. Require explicit scenario analysis of assumption failure regardless of how 'obvious' it appears. Test whether removal or replacement is operationally feasible.

Status	Location	Function	Example	Audit Considerations	Key Audit Steps
Sacred	Structural	Causal	Pre-GFC CDO model embedding the assumption that geographic diversification causes risk reduction in mortgage pools, treated as unchallengeable within the securitisation industry.	Very High. Sacred causal assumption in model architecture; when the causal relationship fails, the entire model produces systematically wrong outputs.	Independent evidence review; challenge whether causal link is empirical or normative; scenario test causal reversal.
Sacred	Structural	Translational	Regulatory capital model translating ratings agency assessments into risk weights using the Standardised Approach, where the translation is mandated by regulation and treated as beyond challenge.	High. Regulatory mandate creates perceived immunity from challenge; risk when ratings prove unreliable as during the GFC.	Distinguish regulatory mandate vs. regulatory permission; assess evidence base; review rating methodology changes.
Sacred	Structural	Transformational	Efficient Market Hypothesis embedded in portfolio optimisation models, treating market prices as always reflecting available information and governing price adjustment dynamics.	High. Sacred transformational assumption in model structure; when market efficiency breaks down (bubbles, crashes), model dynamics are wrong.	Scenario analysis of market inefficiency episodes; test model performance during known anomalies.
Sacred	Structural	Terminal	Too-big-to-fail assumption built into counterparty credit risk models, where systemically important institutions are assumed to never reach default as a terminal state.	Very High. Sacred terminal assumption prevents recognition of catastrophic risk; Lehman Brothers' failure demonstrated this can fail.	Scenario test SIFI default; assess resolution regime credibility; review model behaviour at terminal boundary.
Sacred	Contextual	Baseline	Assumption that central bank independence is permanent and monetary policy will always be conducted by technocratic institutions, serving as the baseline for all interest rate modelling.	High. Widely held but could be challenged by political developments; failure would invalidate baseline rate expectations.	Monitor political risks to central bank independence; scenario test alternative monetary policy regimes.

Status	Location	Function	Example	Audit Considerations	Key Audit Steps
Sacred	Contextual	Causal	Assumption that free capital movement between major economies will persist, causing global risk pricing to remain interconnected, treated as foundational in international finance.	High. Decades of globalisation make this feel permanent; capital controls would break causal assumptions about risk transmission.	Scenario test capital control imposition; assess hedging effectiveness under restricted capital flows.
Sacred	Contextual	Translational	Assumption that LIBOR/SOFR benchmark rates accurately translate the cost of interbank lending into derivative pricing, treated as a definitional truth within market infrastructure.	High. Benchmark manipulation scandals (LIBOR) demonstrated this can fail; LIBOR transition required fundamental repricing.	Review benchmark governance; assess fallback provisions; test sensitivity to benchmark methodology changes.
Sacred	Contextual	Transformational	Assumption that European monetary union is irreversible and the euro cannot fragment, governing how sovereign spreads evolve within the Eurozone.	High. The Eurozone crisis tested this; redenomination risk would fundamentally alter spread dynamics.	Scenario test Eurozone fragmentation; assess redenomination risk in cross-border exposures.
Sacred	Contextual	Terminal	Assumption that deposit insurance schemes will always be honoured by sovereigns, serving as the ultimate terminal safety net preventing systemic bank runs.	High. Generally credible but tested during sovereign crises (Cyprus 2013 bail-in); failure would be catastrophic.	Scenario test sovereign fiscal limits; review historical bail-in precedents; assess systemic contagion paths.
Sacred	Behavioural	Baseline	Assumption that market participants are fundamentally rational and profit-maximising as the behavioural baseline, underpinning equilibrium models across finance.	Medium-High. Foundational in financial economics but repeatedly challenged by behavioural finance evidence.	Review model dependence on rationality assumption; test with bounded-rationality alternatives; assess sensitivity.
Sacred	Behavioural	Causal	Assumption that higher interest rates cause savers to save more and borrowers to borrow less, treated as a bedrock principle of monetary policy transmission.	Medium-High. Generally directionally correct but magnitude is uncertain and can be overridden by wealth effects or expectations.	Backtest elasticities across rate environments; test for wealth effect offsets; assess heterogeneity across segments.

Status	Location	Function	Example	Audit Considerations	Key Audit Steps
Sacred	Behavioural	Translational	Assumption that credit ratings reliably translate into counterparty default behaviour, treated as axiomatic in credit risk management.	High. GFC demonstrated systematic optimism in ratings for structured products; ratings may lag reality.	Backtest rating-default relationship; assess rating timeliness; compare with market-implied default probabilities.
Sacred	Behavioural	Transformational	Assumption that central bank forward guidance reliably shapes market participant expectations and behaviour over time, treated as an established monetary policy tool.	Medium-High. Effectiveness varies; forward guidance can lose credibility if perceived as unreliable.	Assess historical guidance effectiveness; test model sensitivity to guidance credibility shocks.
Sacred	Behavioural	Terminal	Assumption that lender-of-last-resort facilities will prevent solvent banks from reaching terminal liquidity failure, treated as a foundational safety net assumption.	High. Generally credible but access conditions, stigma effects, and political constraints can limit effectiveness.	Scenario test LOLR access restrictions; assess stigma effects; review historical instances of delayed intervention.
Gone	Structural	Baseline	Pre-digitisation fraud detection model using branch transaction patterns as the baseline for detecting anomalies, now obsolete given the shift to online and mobile banking.	High. Outdated structural baseline; model may fail to detect new fraud typologies entirely.	Model inventory review; compare detection rates for digital vs. physical channels; assess update timeline.
Gone	Structural	Causal	Operational risk model assuming physical proximity causes correlated operational losses (e.g., regional natural disasters affecting nearby branches), now obsolete given cyber risk concentration across geographies.	High. Gone causal structure; cyber risk operates through entirely different propagation mechanisms.	Review loss data composition; compare physical vs. cyber loss drivers; assess structural model relevance.
Gone	Structural	Translational	Market risk model translating trading book positions into risk metrics using pre-FRTB methodologies that no longer reflect current regulatory expectations or market structure.	High. Outdated translation methodology; regulatory divergence creates compliance risk.	Gap analysis against current regulatory requirements; assess implementation timeline for updated methodology.

Status	Location	Function	Example	Audit Considerations	Key Audit Steps
Gone	Structural	Transformational	Yield curve model calibrated to a low-interest-rate regime, structurally embedding near-zero rate dynamics that no longer reflect current monetary policy conditions.	High. Structural dynamics calibrated to a gone regime; model behaviour may be unreliable in the current environment.	Recalibrate to include current rate regime; backtest across historical rate environments; test extrapolation behaviour.
Gone	Structural	Terminal	Resolution planning model assuming traditional bank wind-down procedures without accounting for digital asset complexities, fintech interdependencies, or crypto exposures.	Medium-High. Terminal conditions designed for a simpler era; operational executability under current complexity is uncertain.	Review resolution plan against current business model; stress test digital asset liquidation feasibility.
Gone	Contextual	Baseline	Liquidity model using pre-2008 interbank market conditions as the contextual baseline, assuming deep and liquid unsecured interbank lending markets that no longer exist in the same form.	High. Gone contextual baseline; post-GFC interbank markets are fundamentally different (secured vs. unsecured).	Compare baseline against current market structure; assess secured vs. unsecured funding mix; update for post-GFC reality.
Gone	Contextual	Causal	Post-digitisation assumption that high-quality collateral ensures funding access, where the causal link between collateral quality and liquidity held for decades but stopped holding as market structure evolved.	TOXIC COMBINATION: Gone-Contextual-Causal (Regime-Change Combination). The assumption once described reality but no longer does, it defines the operating environment, and its causal chain drives model behaviour. Every causal chain in the model is potentially miswired.	Environmental scanning combined with explicit trigger design. Define observable indicators that signal the assumption has become gone. Test whether causal relationships still hold under current conditions.
Gone	Contextual	Translational	LIBOR-based derivative pricing model translating funding costs into fair values using a benchmark rate that has been discontinued and replaced by SOFR/SONIA.	High. Gone translation mechanism; continued use creates valuation and hedging mismatches.	Review LIBOR-to-RFR transition completeness; assess residual LIBOR exposures; verify fallback provisions.

Status	Location	Function	Example	Audit Considerations	Key Audit Steps
Gone	Contextual	Transformational	Retail banking model assuming customer switching behaviour evolves slowly based on branch proximity, now obsolete given instant digital account opening and comparison platforms.	Medium-High. Gone contextual dynamics; digital switching has fundamentally accelerated competitive dynamics.	Update switching cost assumptions; review digital competitor landscape; test accelerated attrition scenarios.
Gone	Contextual	Terminal	Stress testing model assuming a maximum 5% retail deposit outflow over 30 days as the terminal stress scenario, calibrated before SVB demonstrated 20%+ single-day withdrawals.	Very High. Gone terminal boundary; post-SVB reality demonstrates far more severe and rapid outflows are possible.	Recalibrate terminal scenarios to post-SVB evidence; test 1-day, 1-week, and 30-day speed scenarios.
Gone	Behavioural	Baseline	Customer complaint model using pre-social-media complaint volumes as the behavioural baseline, failing to capture the amplification effect of online platforms.	Medium. Gone behavioural baseline; social media has fundamentally changed complaint behaviour and reputational contagion speed.	Update baseline to include social media metrics; assess reputational risk amplification factors.
Gone	Behavioural	Causal	Branch footprint model assuming physical branch presence causes customer loyalty, now challenged by digital-first banks demonstrating loyalty through app experience and service quality.	Medium-High. Gone causal relationship between physical presence and loyalty; digital experience is now the dominant driver.	Analyse loyalty drivers by channel; compare digital vs. branch customer retention; update causal framework.
Gone	Behavioural	Translational	Customer segmentation model translating demographic characteristics into product preferences using pre-fintech-era purchasing patterns that no longer reflect current consumer behaviour.	Medium. Gone translation of demographics to preferences; fintech has disrupted traditional product adoption patterns.	Backtest demographic-preference mappings; assess fintech adoption patterns by segment; update translation logic.

Status	Location	Function	Example	Audit Considerations	Key Audit Steps
Gone	Behavioural	Transformational	Market panic model assuming information spreads through traditional media channels at newspaper/TV speed, governing how panic transforms from rumour to action over days, not minutes.	High. Gone behavioural dynamics; social media has compressed information propagation and herding behaviour to minutes.	Recalibrate propagation speed assumptions; test social media contagion scenarios; review SVB precedent for speed.
Gone	Behavioural	Terminal	Bank run model assuming physical branch capacity constrains maximum daily withdrawal rates, creating a natural terminal limit on deposit flight speed.	High. Gone terminal constraint; online banking eliminates the physical bottleneck, as SVB demonstrated.	Remove physical capacity constraints from terminal conditions; model digital withdrawal capacity; test concurrent online run scenarios.
Speculative	Structural	Baseline	AI risk model using speculative assumptions about AGI emergence timelines as the structural baseline for technology risk assessment, where no historical precedent exists.	High. Speculative baseline embedded in model architecture; highly uncertain foundation for all downstream analysis.	Label explicitly as speculative; develop wide scenario range; ensure senior governance approval; review regularly.
Speculative	Structural	Causal	Pandemic risk model assuming a speculative causal relationship between zoonotic spillover frequency and deforestation rates, embedded in the model's structural framework.	High. Speculative causal logic in model structure; if the relationship is wrong, the entire risk assessment is compromised.	Review evidence base; compare with alternative causal theories; ensure model outputs labelled as indicative.
Speculative	Structural	Translational	Climate risk model translating carbon emission scenarios into physical risk impacts using speculative damage functions embedded in the model's structural architecture.	High. Speculative translation in model structure; damage functions are highly uncertain and scientifically debated.	Compare multiple damage function specifications; assess sensitivity to translation methodology; disclose uncertainty.

Status	Location	Function	Example	Audit Considerations	Key Audit Steps
Speculative	Structural	Transformational	Climate risk model embedding speculative carbon price trajectories into structural economic relationships that transform asset valuations over multi-decade horizons, compounding initial uncertainty.	TOXIC COMBINATION: Speculative-Structural-Transformational (Compounding-Error Combination). A guess embedded in model architecture governs how states evolve over time. Small errors in initial speculation compound through dynamics, producing outputs diverging progressively from reality.	Dynamic path testing with wide uncertainty bands. Explicit acknowledgement that model outputs at longer horizons are indicative, not predictive. Regular recalibration as new data emerges.
Speculative	Structural	Terminal	Catastrophic risk model assuming speculative tipping points (e.g., ice sheet collapse thresholds) as terminal conditions built into the model's structural framework.	High. Speculative terminal boundaries in model architecture; tipping point locations are inherently uncertain.	Use wide uncertainty ranges for terminal conditions; scenario test multiple tipping points; disclose as speculative.
Speculative	Contextual	Baseline	Geopolitical risk model using a speculative assumption about the future of US-China trade relations as the baseline context for supply chain risk assessment.	Medium-High. Speculative contextual baseline; geopolitical outcomes are inherently unpredictable.	Develop multiple baseline scenarios; benchmark against think-tank forecasts; ensure regular review cycles.
Speculative	Contextual	Causal	Digital currency adoption model speculating that CBDC introduction will cause a specific percentage decline in commercial bank deposits within the operating context.	Medium-High. Speculative causal assumption about an unprecedented contextual change; no historical precedent.	Benchmark against early CBDC adopter data; develop multiple adoption scenarios; monitor pilot programme results.
Speculative	Contextual	Translational	ESG scoring model translating speculative regulatory stringency scenarios into company-level compliance cost estimates within an uncertain future regulatory context.	Medium. Speculative contextual translation; regulatory outcomes are uncertain and jurisdiction-dependent.	Monitor regulatory developments; compare with peer institution estimates; develop regulatory scenario ranges.

Status	Location	Function	Example	Audit Considerations	Key Audit Steps
Speculative	Contextual	Transformational	Long-term energy transition model speculating on how renewable energy cost curves will transform the operating context for fossil fuel asset valuations over 30 years.	Medium-High. Speculative contextual transformation over long horizons; technology cost curves are inherently uncertain.	Compare with multiple technology forecasting sources; test sensitivity to learning curve assumptions; regular recalibration.
Speculative	Contextual	Terminal	Stranded asset model speculating on a future regulatory date by which fossil fuel assets become unlicenseable, creating a terminal condition for carbon-intensive portfolios.	Medium-High. Speculative terminal condition in regulatory context; political and technological uncertainties make timing highly uncertain.	Develop multiple terminal date scenarios; assess portfolio sensitivity to timing; monitor policy developments.
Speculative	Behavioural	Baseline	Consumer adoption model using speculative assumptions about baseline customer willingness to share biometric data for financial services in the absence of historical adoption data.	Medium. Speculative behavioural baseline; consumer privacy attitudes are evolving and culturally variable.	Benchmark against early adopter data; survey-based validation; develop optimistic and pessimistic baselines.
Speculative	Behavioural	Causal	AI-driven trading model speculating that algorithmic herding will cause flash crash frequency to increase by a specific rate as AI adoption increases in trading.	Medium-High. Speculative causal assumption about emergent behaviour in complex adaptive systems.	Review evidence from past flash crashes; monitor AI trading adoption metrics; develop multiple frequency scenarios.
Speculative	Behavioural	Translational	Gen-Z banking model translating speculative assumptions about generational financial attitudes into product demand forecasts for the next decade.	Medium. Speculative behavioural translation; generational preference patterns may not persist or may evolve unpredictably.	Track emerging data on generational banking behaviour; compare with historical generational shifts; regular recalibration.
Speculative	Behavioural	Transformational	Social media contagion model speculating on how misinformation about a bank's solvency would transform depositor behaviour over a 48-hour crisis window.	Medium-High. Speculative behavioural transformation; social media dynamics during crises are highly uncertain and non-linear.	Review SVB and similar precedents; simulate contagion scenarios; test institutional communication response plans.

Status	Location	Function	Example	Audit Considerations	Key Audit Steps
Speculative	Behavioural	Terminal	Cybersecurity model speculating that a successful quantum computing attack on banking encryption would create a terminal loss of customer trust, permanently changing banking behaviour.	Medium. Speculative terminal behavioural state; quantum computing timeline and impact are highly uncertain.	Monitor quantum computing developments; assess encryption migration plans; scenario test trust recovery timelines.

Appendix II - Common Cognitive Biases

Cognitive Bias	Definition	Example in Financial Modelling / Audit	Evolutionary Explanation	Key Things for Auditors to Look Out For
Availability Heuristic	Tendency to overweight information that comes easily to mind, such as recent events, vivid examples, or familiar techniques, when making judgements.	Modellers gravitate towards normal distributions and standard software techniques because these are readily available in educational materials and toolkits, even when alternative distributions would be more appropriate for the data.	In ancestral environments, easily recalled events (predator attacks, food sources) were often the most relevant to survival. Quick recall served as a proxy for frequency and importance.	Check whether model choices reflect genuine suitability or merely familiarity. Ask: 'Why was this technique chosen over alternatives?' Look for over-reliance on standard textbook approaches without justification.
Anchoring Bias	Tendency to rely too heavily on the first piece of information encountered (the 'anchor') when making subsequent judgements, with insufficient adjustment away from that anchor.	Once a simplifying assumption is adopted in an initial model version, subsequent iterations retain it without re-examination because it was 'the original assumption.' Baseline assumptions are heavily influenced by anchoring to recent economic conditions.	In environments with limited information, using an initial reference point and adjusting from it was an efficient decision-making strategy. The problem arises when we fail to adjust sufficiently.	Review whether baseline assumptions extrapolate recent conditions rather than reflecting structural change. Check whether model parameters have been updated or merely carried forward from prior versions.
Confirmation Bias	Tendency to search for, interpret, and recall information in a way that confirms pre-existing beliefs or hypotheses, while giving less attention to contradictory evidence.	Practitioners preferentially notice evidence supporting the 'obvious' risk-return relationship and discount anomalies as temporary. Subsequent model validation focuses on parameters within the existing structure rather than questioning the structure itself.	Seeking confirmation of existing beliefs conserved cognitive energy and supported group cohesion in tribal settings. Changing beliefs was costly and potentially dangerous to social standing.	Look for model validation that only tests whether parameters are correct within the existing framework without challenging the framework itself. Ask whether contradictory evidence has been actively sought and considered.

Cognitive Bias	Definition	Example in Financial Modelling / Audit	Evolutionary Explanation	Key Things for Auditors to Look Out For
Conformity Bias	Tendency to align one's beliefs, attitudes, and behaviours with those of the group, even when the group may be wrong. Also known as social proof or herd mentality in judgement.	There is social and professional pressure to align with consensus on 'obvious' assumptions. Challenging a widely held assumption may be perceived as contrarian without purpose, discouraging necessary scepticism.	In ancestral groups, conforming to group norms was essential for survival. Dissent could lead to exclusion from the group, which was often fatal in the wild.	Assess whether assumption challenges are genuinely considered or suppressed by group dynamics. Look for evidence that dissenting views are documented and addressed rather than rejected.
Authority Bias	Tendency to attribute greater accuracy to the opinion of an authority figure and to be influenced by that opinion even when the evidence is ambiguous or contradictory.	Sacred assumptions persist partly because they are endorsed by regulators, rating agencies, academic authorities, or industry-wide practice. Challenging regulatory assumptions is perceived as challenging the regulator itself.	In tribal settings, following the judgement of experienced leaders (who had survived to become authorities) was a reliable survival strategy. Deference to authority was adaptive.	Distinguish between assumptions mandated by regulation and those merely permitted by it. Ask: 'Is this required, or is it a choice that happens to align with authority?' Review whether regulatory assumptions are empirically tested.
Groupthink	A psychological phenomenon where the desire for conformity and harmony within a group leads to irrational or dysfunctional decision-making, with dissenting viewpoints suppressed or ignored.	Sacred assumptions are sustained by groupthink; dissenting views on risk-free sovereign debt, uncorrelated mortgage defaults, or perpetual liquidity were suppressed because they threatened the consensus that the entire industry relied upon.	Group cohesion was critical for survival in ancestral environments. Mechanisms that promoted agreement and suppressed dissent strengthened group bonds, even at the cost of decision quality.	Review governance forum minutes for evidence of genuine debate vs. rubber-stamping. Multiple sacred assumptions in a single model may indicate groupthink. Look for absence of challenge documentation.
Status Quo Bias	Preference for the current state of affairs, where changes from the baseline are perceived as losses. The default option is favoured even when alternatives may be objectively better.	Gone assumptions persist because changing established models is complex and costly. Legacy operational risk models built for physical-era risks continue despite digital transformation, because the status quo feels safer than change.	In stable ancestral environments, maintaining existing successful strategies was generally safer than experimenting. The cost of change (potential failure) was high relative to marginal improvement.	Identify models not substantially updated since known structural changes. Ask: 'When was this model last fundamentally reviewed?' Challenge assumptions that persist solely because changing them is difficult.

Cognitive Bias	Definition	Example in Financial Modelling / Audit	Evolutionary Explanation	Key Things for Auditors to Look Out For
Sunk Cost Fallacy	Tendency to continue investing in a decision, project, or course of action because of previously invested resources (time, money, effort), rather than evaluating current and future value.	Resources invested in building models based on old assumptions are used to justify continued use of those models, even when the underlying assumptions are known to be gone or inappropriate.	In ancestral environments, abandoning a partially completed shelter or hunt meant losing tangible survival resources. Persisting with existing investments was often the safer bet.	Ask whether continued use of a model is justified by its current purpose or by replacement costs. Look for cost-of-change arguments being used to avoid necessary model updates.
Recency Bias	Tendency to overweight recent events or data relative to earlier observations, leading to extrapolation of current conditions and underweighting of historical patterns that may recur.	Modellers overweight recent stable conditions and underweight historical periods of disruption when calibrating models. Contextual assumptions are shaped by the expectation that the future will resemble the recent past. May delay recognition that an assumption is 'gone'.	In changing environments, recent information was typically more relevant to immediate survival than distant memories. Prioritising recent data was an efficient information processing strategy.	Check model calibration periods: do they include stress episodes? Ask whether recent benign conditions are assumed to continue. Look for scenarios anchored to the most recent crisis rather than exploring different failure modes.
Normalcy Bias	Tendency to underestimate the possibility and impact of a disaster or major disruption, assuming that the future will continue to operate as it normally has.	Contextual assumptions are shaped by normalcy bias, expecting future conditions to resemble the recent stable past. Baseline assumptions often extrapolate current conditions rather than reflecting potential structural change. Terminal assumptions may underestimate the probability of reaching boundary conditions.	In ancestral environments, the status quo persisting was the most common outcome. Catastrophes were rare, and cognitive resources spent worrying about unlikely events reduced capacity for daily survival tasks.	Test whether models and scenarios adequately consider low-probability, high-impact events. Ask: 'What would happen if this assumption fails completely?' Look for stress tests that are insufficiently severe.

Cognitive Bias	Definition	Example in Financial Modelling / Audit	Evolutionary Explanation	Key Things for Auditors to Look Out For
Overconfidence Bias	Tendency to overestimate the accuracy of one's own knowledge, predictions, and ability to assess risk. Manifests as overly narrow confidence intervals and excessive certainty in forecasts.	Speculative assumptions are particularly vulnerable: modellers overestimate the accuracy of their predictions about uncertain futures like carbon prices, AI productivity gains, or pandemic impacts. Point estimates are presented without adequate uncertainty ranges.	In ancestral environments, confidence was adaptive: hesitant individuals were less likely to take necessary risks (hunting, exploration) and confident individuals attracted more social support and mating opportunities.	Review whether uncertainty ranges are calibrated to historical forecast accuracy. Ask whether speculative assumptions are labelled as such. Look for point estimates presented without confidence intervals or scenario ranges.
Optimism Bias / Pessimism Bias	Systematic tendency to overestimate (optimism) or underestimate (pessimism) the likelihood of positive vs. negative events. Often influenced by the modeller's incentives and disposition.	Baseline assumptions may embed optimistic economic projections; strategic plans may assume current market share is sustainable when competition is intensifying. Conversely, overly pessimistic stress scenarios may be used to justify excessive capital buffers.	Optimism bias in particular is well-documented evolutionarily: moderate optimism motivated action, exploration, and risk-taking that were necessary for survival. Pessimism served as a protective mechanism in dangerous situations.	Compare baselines against external consensus; assess whether past baselines have been consistently biased in one direction. Check whether modeller incentives align with optimistic or pessimistic outcomes.
Projection Bias	Tendency to assume that other people share the same beliefs, values, and behaviours as oneself, or to project one's current emotional state onto future decisions.	Modellers assuming counterparties or customers will behave as they would under similar conditions. Behavioural assumptions about rational exercise of early termination options may reflect the modeller's own rationality rather than actual counterparty behaviour.	In small ancestral groups, members shared similar experiences, values, and constraints. Assuming others thought and acted similarly was a reasonable approximation in homogeneous tribal settings.	Challenge whether behavioural assumptions are based on observed data or projected expectations. Ask: 'How do we know customers/counterparties will actually behave this way?' Seek empirical behavioural evidence.

Cognitive Bias	Definition	Example in Financial Modelling / Audit	Evolutionary Explanation	Key Things for Auditors to Look Out For
Fundamental Attribution Error	Tendency to attribute others' behaviour to their character or disposition rather than to situational factors, while attributing one's own behaviour to situational factors.	When sacred assumptions fail, the failure is attributed to external shocks (situational) rather than flaws in the assumption itself (dispositional). Conversely, model failures at other institutions are attributed to poor judgement rather than shared systemic vulnerabilities.	In small groups, judging others' character from behaviour was a useful shortcut for predicting future interactions. Situational awareness for one's own actions helped maintain flexible self-assessment.	Look for post-mortem analyses that blame external factors without examining whether assumptions were inherently flawed. Ask: 'If the same shock hit us, would our assumptions have held?'
Illusory Correlation	Tendency to perceive a relationship between two variables where none exists, or to overestimate the strength of a real but weak relationship.	Causal assumptions may be based on spurious correlations in historical data. Two variables driven by a common third factor may be assumed to have a direct causal link, leading to model misspecification.	Pattern recognition was a critical survival skill: recognising correlations between events (e.g., certain weather patterns and predator behaviour) helped ancestors make protective decisions, even at the cost of some false positives.	Demand evidence for causal links beyond correlation. Ask: 'Could a third factor explain this relationship?' Look for causal assumptions based solely on observed co-movement without theoretical foundation.
Post Hoc Ergo Propter Hoc	Fallacy of assuming that because event B followed event A, event A caused event B. Confuses temporal sequence with causation.	A model assumes that because interest rate increases have historically been followed by credit spread widening, rate increases cause spread widening. The actual mechanism may involve inflation expectations, risk appetite, or other intermediating factors.	In ancestral environments, temporal sequences were often the best available evidence for causation (eating a plant followed by illness suggested the plant was poisonous). Speed of learning outweighed precision.	For each causal assumption, ask: 'What is the mechanism?' Test whether the assumed causal relationship holds across different time periods and regimes. Look for intermediating variables.

Cognitive Bias	Definition	Example in Financial Modelling / Audit	Evolutionary Explanation	Key Things for Auditors to Look Out For
Representativeness Heuristic	Tendency to judge the probability of an event by how much it resembles a typical case or prototype, rather than by using actual statistical reasoning.	Translational assumptions assume that a metric in one domain accurately represents a concept in another (e.g., CDS spreads 'represent' default risk, or credit ratings 'represent' creditworthiness), when the metric may embed additional factors like liquidity premia and sentiment.	In ancestral environments, judging by similarity (this berry looks like the safe ones) was an efficient classification strategy. The cost of occasional misclassification was outweighed by the speed of decision-making.	Question whether translated metrics truly represent what they are assumed to represent. Ask: 'What else does this metric capture beyond what we're using it for?' Assess information loss in translations.
Linearity Bias	Tendency to assume linear relationships and proportional responses when actual relationships may be non-linear, asymmetric, or subject to threshold effects.	Transformational assumptions often assume parallel yield curve shifts, proportional inflation impacts, or linear transmission of policy changes. In reality, many financial relationships are non-linear with tipping points and regime-dependent responses.	Linear thinking was sufficient for most ancestral decision-making contexts, where relationships between effort and reward, distance and time, etc., were approximately linear within relevant ranges.	Test for non-linear responses by varying assumption magnitude. Ask: 'Does a 2x change in input produce a 2x change in output?' Look for threshold effects and regime-dependent behaviour.
Extrapolation Bias	Tendency to project current trends indefinitely into the future without considering structural breaks, mean reversion, or regime changes.	Transformational assumptions governing path-dependent processes may extrapolate recent dynamics (e.g., low volatility, stable correlations) into the future. Terminal growth rate assumptions in DCF models may extrapolate recent growth without considering competitive dynamics.	In stable ancestral environments, extrapolating recent trends (seasonal patterns, resource availability) was a reliable forecasting method. The concept of structural breaks was less relevant.	Check whether trend extrapolation accounts for mean reversion, structural breaks, and competitive dynamics. Ask: 'What would cause this trend to stop or reverse?' Review DCF terminal assumptions critically.

Cognitive Bias	Definition	Example in Financial Modelling / Audit	Evolutionary Explanation	Key Things for Auditors to Look Out For
Bias Blind Spot	Tendency to recognise the influence of biases on other people's judgements but fail to recognise the same biases in one's own thinking. A well-documented meta-bias.	Model developers acknowledge biases exist in general but believe their own models and assumptions are objective. 'Everybody is biased' (as the whitepaper notes), yet individuals consistently rate their own judgement as less biased than others'.	Self-serving assessment of one's own objectivity supported confidence and social status in ancestral groups. Recognising one's own biases would have been demotivating and socially costly.	Use structured taxonomies and frameworks (like the assumption classification in this whitepaper) to systematically identify biases rather than relying on individual self-awareness. Ensure independent review by parties without stake in the model.
Dunning-Kruger Effect	Cognitive bias where individuals with limited competence in a domain overestimate their own ability, while highly competent individuals tend to underestimate theirs.	A business-line model owner with limited statistical training may dismiss model risk concerns, believing the model is simpler than it actually is. Conversely, experienced quants may understate the reliability of well-calibrated models due to awareness of edge cases.	In ancestral environments, attempting tasks beyond one's skill level (overconfidence) had survival value through trial and learning. The mechanism for accurate self-assessment was less developed because feedback was often immediate and physical.	Assess whether model governance includes appropriately qualified reviewers. Look for mismatch between model complexity and reviewer expertise. Ensure challenge comes from individuals with sufficient technical depth.
Endowment Effect	Tendency to overvalue things simply because one owns them or has created them. People demand more to give up an object than they would pay to acquire it.	Model developers overvalue their own models and resist replacing them, even when evidence suggests alternatives would be more appropriate. 'Our model' becomes more valuable than 'a better model' because of emotional ownership.	In ancestral environments, objects in one's possession had concrete survival value and were costly to replace. Valuing what you have over what you could acquire reduced risky exchanges.	Assess whether model retention decisions are based on fitness for purpose or on attachment. Ask: 'If we were building this from scratch today, would we make the same choices?' Look for defensiveness about model limitations.

Cognitive Bias	Definition	Example in Financial Modelling / Audit	Evolutionary Explanation	Key Things for Auditors to Look Out For
In-Group Bias	Tendency to favour members of one's own group over outsiders, extending to preferring ideas, methods, and assumptions originating from within the group.	Model validation teams within the same institution may give preferential treatment to internally developed models vs. vendor models, or vice versa. Audit challenges from external parties may be dismissed as 'not understanding our business.'	Strong in-group preference was critical for ancestral survival: cooperating with and trusting one's own tribe members, while being wary of outsiders, increased group survival probability.	Assess whether external challenge (auditors, validators, regulators) is genuinely engaged with or dismissed. Look for patterns where internal models receive less scrutiny than external ones, or where external expertise is systematically discounted.

References

1. Seminal Academic Papers and Books

Financial Modelling and Pricing Foundations

Black, F. and Scholes, M. (1973). "The Pricing of Options and Corporate Liabilities." The Journal of Political Economy, 81(3), pp. 637–654.

Merton, R.C. (1973). "Theory of Rational Option Pricing." The Bell Journal of Economics and Management Science, 4(1), pp. 141–183.

Merton, R.C. (1974). "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates." The Journal of Finance, 29(2), pp. 449–470.

Li, D.X. (2000). "On Default Correlation: A Copula Function Approach." The Journal of Fixed Income, 9(4), pp. 43–54.

Behavioural Economics and Cognitive Bias

Kahneman, D. and Tversky, A. (1979). "Prospect Theory: An Analysis of Decision under Risk." Econometrica, 47(2), pp. 263–292.

Kahneman, D. (2011). Thinking, Fast and Slow. New York: Farrar, Straus and Giroux.

Pronin, E., Lin, D.Y. and Ross, L. (2002). "The Bias Blind Spot: Perceptions of Bias in Self Versus Others." Personality and Social Psychology Bulletin, 28(3), pp. 369–381.

Knight, F.H. (1921). Risk, Uncertainty, and Profit. Boston and New York: Houghton Mifflin Company.

Keynes, J.M. (1921). A Treatise on Probability. London: Macmillan and Co.

Keynes, J.M. (1936). The General Theory of Employment, Interest and Money. London: Macmillan and Co.

Read, C. (1906). Logic, Deductive and Inductive. London: Alexander Moring.

Taleb, N.N. (2001). Fooled by Randomness: The Hidden Role of Chance in Life and in the Markets. New York: TEXERE.

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

Taleb, N.N. (2020). Statistical Consequences of Fat Tails: Real World Preasymptotics, Epistemology, and Applications. STEM Academic Press.

Model Risk

Derman, E. (1996). "Model Risk." Goldman Sachs Quantitative Strategies Research Notes.

Derman, E. (2011). Models.Behaving.Badly: Why Confusing Illusion with Reality Can Lead to Disaster, on Wall Street and in Life. New York: Free Press.

Rebonato, R. (2007). Plight of the Fortune Tellers: Why We Need to Manage Financial Risk Differently. Princeton, NJ: Princeton University Press.

Danielsson, J., James, K., Valenzuela, M. and Zer, I. (2014). "Model Risk of Risk Models." Finance and Economics Discussion Series No. 2014-34, Board of Governors of the Federal Reserve System.

Brown, J.A., McGourty, B. and Schuermann, T. (2015). "Model Risk and the Great Financial Crisis: The Rise of Modern Model Risk Management." Wharton Financial Institutions Center Working Paper No. 15-01.

Green, T.C. and Figlewski, S. (1999). "Market Risk and Model Risk for a Financial Institution Writing Options." Journal of Finance, 54(4), pp. 1465–1499.

Post-GFC Copula Model Criticism

Salmon, F. (2009). "Recipe for Disaster: The Formula That Killed Wall Street." Wired, 17(3).

MacKenzie, D. and Spears, T. (2014). "'The Formula That Killed Wall Street': The Gaussian Copula and Modelling Practices in Investment Banking." Social Studies of Science, 44(3), pp. 393–417.

MacKenzie, D. and Spears, T. (2014). "'A Device for Being Able to Book P&L': The Organizational Embedding of the Gaussian Copula." Social Studies of Science, 44(3), pp. 418–440.

Donnelly, C. and Embrechts, P. (2010). "The Devil is in the Tails: Actuarial Mathematics and the Subprime Mortgage Crisis." ASTIN Bulletin: The Journal of the IAA, 40(1), pp. 1–33.

MacKenzie, D. (2011). "The Credit Crisis as a Problem in the Sociology of Knowledge." American Journal of Sociology, 116(6), pp. 1778–1841.

FVA Debate Literature

Hull, J.C. and White, A. (2012). "The FVA Debate." Risk, 25th Anniversary Edition, pp. 83–85.

Hull, J.C. and White, A. (2014). "Valuing Derivatives: Funding Value Adjustments and Fair Value." Financial Analysts Journal, 70(3), pp. 46–56.

Piterbarg, V. (2010). "Funding beyond Discounting: Collateral Agreements and Derivatives Pricing." Risk Magazine, 23(2), pp. 97–102.

Burgard, C. and Kjaer, M. (2011). "Partial Differential Equation Representations of Derivatives with Bilateral Counterparty Risk and Funding Costs." The Journal of Credit Risk, 7(3), pp. 1–19.

Andersen, L., Duffie, D. and Song, Y. (2019). "Funding Value Adjustments." Journal of Finance, 74, pp. 145–192.

2. Regulatory and Supervisory References

US Federal Model Risk Management

Board of Governors of the Federal Reserve System and Office of the Comptroller of the Currency (2011). Supervisory Guidance on Model Risk Management. SR 11-7 / OCC Bulletin 2011-12.

UK Prudential Regulation

Prudential Regulation Authority, Bank of England (2023). SS1/23 - Model Risk Management Principles for Banks. Supervisory Statement.

European Banking Authority

European Banking Authority (2017). Guidelines on Credit Institutions' Credit Risk Management Practices and Accounting for Expected Credit Losses. EBA/GL/2017/06.

European Banking Authority (2022). Guidelines on the Management of Interest Rate Risk Arising from Non-Trading Book Activities and Credit Spread Risk Arising from Non-Trading Book Activities. EBA/GL/2022/14.

IFRS Accounting Standards

International Accounting Standards Board (2014). IFRS 9 Financial Instruments.

International Accounting Standards Board (2011). IFRS 13 Fair Value Measurement.

US Accounting Standards

Financial Accounting Standards Board (2006). Statement of Financial Accounting Standards No. 157: Fair Value Measurements.

Basel Committee on Banking Supervision

Basel Committee on Banking Supervision (2006). Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework - Comprehensive Version. BCBS 128.

Basel Committee on Banking Supervision (2013). Principles for Effective Risk Data Aggregation and Risk Reporting. BCBS 239.

Basel Committee on Banking Supervision (2015). Guidance on Credit Risk and Accounting for Expected Credit Losses. BCBS d350.

3. Scenario Probability Assignment: IFRS 9 ECL Requirements and Guidance

IASB and Regulatory Guidance

IASB Staff Paper (March 2024). "Post-implementation Review of IFRS 9 - Impairment: Feedback Analysis - Measuring Expected Credit Losses." Agenda Reference 27A.

ECB Banking Supervision (July 2024). "IFRS 9 Overlays and Model Improvements for Novel Risks."

EBA (November 2023). "IFRS 9 Implementation by EU Institutions: 2023 Monitoring Report."

ECB SSM Supervisory Priorities 2025–2027 (December 2024).

PRA (September 2024). "Thematic Feedback on Accounting for IFRS 9 ECL and Climate Risk."

Academic and Practitioner Criticism

Canals-Cerdá, J.J. (2024). "CECL Implementation and Model Risk in Uncertain Times: An Application to Consumer Finance." Federal Reserve Bank of Philadelphia Working Paper 24-03.

Chawla, G., Forest, L.R. and Aguais, S.D. (2016). "Point-in-Time Loss-Given Default Rates and Exposures at Default Models for IFRS 9/CECL and Stress Testing." Journal of Risk Management in Financial Institutions, 9(3), pp. 249–263.

Chawla, G., Forest, L. and Aguais, S. (2016). "Some Options for Evaluating Significant Deterioration under IFRS 9." The Journal of Risk Model Validation, 10(3), pp. 69–89.

Behn, M. and Couaillier, C. (2023). "Same Same but Different: Credit Risk Provisioning under IFRS 9." ECB Working Paper Series No. 2841.

Buesa, A., Población García, F.J. and Tarancón, J. (2019). "Measuring the Procyclicality of Impairment Accounting Regimes: A Comparison between IFRS 9 and US GAAP." ECB Working Paper Series No. 2347.

GARP Risk Intelligence (March 2023). "When to Change IFRS 9 Scenario Weights for ECL: A Simple Rule."

4. Professional Standards and Risk Management Frameworks

Institute of Internal Auditors

The Institute of Internal Auditors (2024). Global Internal Audit Standards (The Redbook).

The Institute of Internal Auditors (2018). Practice Guide: Auditing Model Risk Management.

The Institute of Internal Auditors (2020). Practice Guide: Auditing Market Risk in Financial Institutions.

The Institute of Internal Auditors (2020). The IIA's Three Lines Model: An Update of the Three Lines of Defense.

COSO Frameworks

Committee of Sponsoring Organizations of the Treadway Commission (2017). Enterprise Risk Management - Integrating with Strategy and Performance.

Committee of Sponsoring Organizations of the Treadway Commission (2013). Internal Control - Integrated Framework.

ISO Risk Management

International Organization for Standardization (2018). ISO 31000:2018 - Risk Management - Guidelines.

5. Layered Beliefs Framework/ Assumptions N-Space Taxonomy

Herfel, Julius (2025). Layered Beliefs Series of Articles. (in particular “LB2: Judgement Day for Expert Judgement”, “LB3: When Models Pick Sides”, “LB4: Beautiful Mess”, “LB8: The Ape in the Algorithm”, and “LB9: Oracles with Error Bars”)

Herfel, Julius (2026). The Big Mistake: How Tech Solved Model Risk by Firing the Modeller.

Herfel, Julius (2026, forthcoming). Overconfidently Into The Abyss.

Disclaimer

This publication is for general information purposes only and does not constitute professional, financial, legal, or investment advice. This whitepaper is not intended as a financial promotion, nor does it constitute an invitation, inducement, or solicitation to engage in any form of investment activity. Interpretations of case law and accounting standards are provided for illustrative purposes only and do not constitute legal or accounting advice. Technical descriptions reflect current practices as of February 2026 and may evolve.

The frameworks and methodologies presented are illustrative and should be adapted to specific institutional contexts and regulatory requirements. The views expressed are personal and do not represent the official position of any organisation. Audit professionals should apply professional scepticism and judgement in all circumstances. This whitepaper provides tools and frameworks to support audit execution but does not replace professional standards, regulatory requirements, or institutional policies.

Herfel Limited and the author of this whitepaper are not liable for any loss, regulatory fines and sanctions, or other detrimental business consequences should macroeconomic conditions, technological shifts, or sudden regime changes render the discussed examples or predictive methodologies obsolete.

Herfel Limited and the author of this whitepaper are not responsible for the accuracy and adequacy of the AI prototype, or the training sets provided. While these are free for use for any reader and have been tested, assumptions may change status, serve a different function or be situated in a different location than envisaged during training. Also, AI may produce results which are wrong. Human oversight and review is required. Lastly, like every model, the prototype was built based on assumptions which might not always be fit for purpose in every circumstance. Any third-party frameworks, regulatory text (such as EBA guidelines, IFRS standards, or PRA Supervisory Statements), or underlying data utilised to train the open-source AI prototype remain the exclusive intellectual property of their respective originating authorities or corporate owners.

Financial institutions deploying the prototype are solely and non-delegably responsible for conducting independent model validation, conceptual soundness checks, and ongoing outcomes analysis in strict accordance with SR 11-7, PRA SS1/23 and similar regulations. The burden of regulatory compliance remains entirely with the deploying institution. The use of the proposed assumption taxonomy, the core audit tools, or the AI prototype does not grant any form of regulatory safe harbor. Institutions remain fully and independently accountable for their internal Model Risk Management (MRM) frameworks and the accuracy of the capital requirements derived from those governed models.