The Register Has No Memory [1]

Pricing the cost of an independence assumption on HS2, across schedule and cost, and a low-cost correction available to any project

Coupled schedule and cost Bayesian networks on HS2: the priced independence gap, with structural, distributional and component-breach tests

Nikhil Dhand

Abstract

This paper presents a low-cost diagnostic for the quantitative risk modelling of megaprojects and prices, on High Speed 2 (HS2), what the most common modelling shortcut costs. In 2013 the programme's quantitative risk model treated its threats as broadly independent, a choice stated explicitly in clause 4.2.3 of its *Cost and Risk Status Report* (Department for Transport and HS2 Ltd, 2013). On a megaproject the principal threats are not independent: adverse ground conditions raise the likelihood of tunnelling delay; tunnelling delay extends the programme; an extended programme carries a standing-army cost; and that cost propagates to inflation and the supply chain. Risk accumulates in one direction, in a way an independence assumption cannot represent.

I reconstruct HS2's Phase 1 risk as two coupled Bayesian networks. The first is a ten-system **schedule** model whose output is the programme slip in days. The second is a twelve-system **cost** model in which one node, Schedule Delay Cost, is derived from that slip; so that schedule risk enters the cost answer through an explicit, documented mechanism rather than by assertion. Each network is run through several configurations, holding the engine and the random seed (4209) fixed, and varying only how much dependence the model is permitted to represent.

On **schedule**, switching the dependencies on, and changing nothing else, raises the P90 programme slip from 3,671 days (about 10.1 years) to 4,535 days (about 12.4 years); roughly 864 days, or 2.4 years, of schedule exposure attributable to model structure alone. The dependency-aware slip falls inside the official reset of a 2036–2043 opening, a slip of ten to seventeen years (HS2 Ltd and Department for Transport, 2026).

On **cost**, the same switch raises the P90 from £67.4bn to £88.6bn: approximately £21.3bn of cost exposure attributable to structure alone, before any new information

[1] How to cite this paper: Dhand, N. (2026). The Register Has No Memory; *PM World Journal*, Vol. XV, Issue VIII, August.

about the project is introduced. A bootstrap of the raw output places the sampling error on that figure at about £0.1bn. The dependency-aware cost chain reaches the floor of the official reset range of £87.7bn to £102.7bn from the coupled register alone. Updating the chain on evidence, with each document admitted only at the cut-off on or after its date, carries it to the band and past it: the public record alone reaches £91.8bn, inside the band, and the system-level magnitudes a sponsor already holds reach £105.4bn, about £2.7bn above its ceiling, with schedule counterparts of 4,671 and 6,187 days. The structure sets the level; disclosure carries the rest, and the distance between the two updated lenses measures the disclosure gap directly. Decomposing each gap, using a chain run with the noisy-OR combination rule disabled, shows that the conditional structure, the bare fact that the risks are connected, carries most of it: 74 per cent of the cost gap and 59 per cent of the schedule gap, with the noisy-OR rule supplying the remainder.

The result is hardened against the obvious objections. A *topology test* (eight runs across reversed, common-cause, sparse and dense variants of each graph) leaves the independence gap intact in every case (£17.6bn to £21.7bn on cost; 833 to 895 days on schedule). A *distributional-shape test* (lognormal magnitudes in place of three-point ranges) moves the P90 by under one per cent. And a *component-breach test*, run on the per-node output, shows the dependence lifting both the expected number of simultaneous breaches of the independent ceilings (from 0.60 to 0.96 among the six cost drivers) and (the property an independence assumption most conceals) its variance (from 0.54 to 0.80), the statistical signature of co-movement.

The implication for a sponsor is direct. The same registers, run twice, would have revealed most of these gaps at the outset, at the cost of one additional simulation each. Every headline figure is recomputed from the raw output by a separate check, and the structural-robustness sweep, the distributional test and the component-breach test are reported in full.

Keywords: *quantitative cost-risk analysis; schedule-risk analysis; Bayesian networks; Monte Carlo simulation; dependence and correlation; schedule-driven cost; reference-class forecasting; megaproject governance; HS2.*

1. Introduction

1.1 Motivation and scope

This paper is not written as a criticism of HS2. It is written because the programme was poorly served by the way its risk was modelled, and because the correction is inexpensive enough that HS2, and comparable projects, can still adopt it. The argument reduces to one quantity and one practice.

The quantity is the cost of assuming that risks are independent. On HS2's own register, with no additional information about the project, that assumption concealed approximately £21.3bn of cost exposure and about 2.4 years of schedule exposure. Any project can estimate its own equivalent figures on the register it already holds, by running that register twice: once in its current configuration, and once with the dependencies between risks switched on. The difference between the two P90 values is the project's *independence error*, expressed in pounds and in days. On a real programme it is not small.

The practice is allowing the baseline to move. A figure set at the start of a project, when the project knows the least it will ever know about itself, and then defended through successive reviews for a decade, drifts away from the truth a little more each year. A figure that is updated as evidence arrives does not. Neither the quantity nor the practice requires bespoke software, and neither requires data that HS2 did not already possess. The register existed; the estimators existed; the public record existed. What was missing in 2013 was a model structure that allowed the risks to influence one another, and the discipline to update the answer as conditions changed.

I have been deliberate about the limits of the demonstration. The method is shown here on a single project whose outcome was already known when the analysis was performed, so it constitutes a worked case rather than a validated forecasting tool. The central finding, however, does not depend on hindsight or on any value I introduced. It depends only on the registers HS2 published and on whether its risks were genuinely independent. The evidence presented here indicates that they were not.

1.2 The independence assumption and the defended baseline

In October 2013 the programme's quantitative risk model presented a register of active threats, each carrying a cost impact and a mitigation plan and described as under continual review. Clause 4.2.3 of the *Cost and Risk Status Report* (Department for Transport and HS2 Ltd, 2013) records the modelling assumption in plain terms: the threats were generally modelled as independent events, with covariance introduced between some related threats only where an analyst deemed it appropriate. On a project in which adverse ground conditions raise the likelihood of tunnelling delay, which extends the programme, which raises cost, which propagates to the supply chain, the model was instructed to treat the principal risks as mutually unrelated.

The single correction that admits a relationship, hand-applied covariance, was added only where an analyst had already noticed a connection; every connection that went unnoticed remained independent. Independence is not a cautious assumption that errs toward safety. It is a precise claim that the hazards cannot arrive together, and on a construction site that claim is false. The Phase 1 forecast has since been reset to between £87.7bn and £102.7bn, with the opening slipping to between 2036 and 2043, and the railway still not in operation (HS2 Ltd and Department for Transport, 2026).

The deeper problem is not specific to HS2 and merits separate statement. A project sets its baseline once, at the outset, which is the moment of least information. That figure is then written into a Treasury submission and defended, review after review, frequently by people who were not present when it was set. As the project proceeds (a contractor struggles, the ground proves worse than the survey indicated, inflation accelerates, a station's scope expands) reality drifts, and because the baseline is frozen, the gap between the defended figure and the project's emerging cost widens steadily. That widening gap is the overrun. It does not originate in the month the cranes stand idle; it accumulates, in plain sight, throughout. This is the documented signature of optimism bias and the planning fallacy in large projects (Kahneman, 2011), and the corrective adopted here (anchoring estimates to the distribution of comparable completed projects rather than to a project-specific narrative) follows the reference-class tradition established for public works (Flyvbjerg, Holm and Buhl, 2002).

1.3 The contribution of this paper

The components of the method are individually well established. The representation of risks as a network of cause-and-effect relationships is the standard apparatus of Bayesian networks (Pearl, 1988; Jensen and Nielsen, 2007), whose application to risk assessment is set out by Fenton and Neil (2018). The treatment of statistical dependence between simulated quantities through induced rank correlation is routine in quantitative risk practice (Iman and Conover, 1982; Vose, 2008), as are its known limitations (Embrechts, McNeil and Straumann, 2002). The use of comparable past projects to bound forecasting optimism is the reference-class method (Flyvbjerg, Holm and Buhl, 2002; Flyvbjerg and Gardner, 2023).

The contribution of this paper is not a new technique but the assembly of these elements into a single diagnostic, applied for the first time to HS2 across both schedule and cost, with the two domains explicitly coupled. Specifically, the paper (i) builds a schedule network and isolates the effect of the independence assumption on the programme slip; (ii) derives, through a documented mechanism, the cost of that slip and feeds it to the cost network as a single node; (iii) isolates and prices the effect of the independence assumption on the cost answer, holding all magnitudes fixed; (iv) updates each network on the public record and then on the sponsor's own observation log, and shows where each lands against the official reset; (v) tests the gap against a topology sweep, a lognormal shape change and a per-node component-breach test; and (vi) audits every headline figure against the raw output. The forward, rolling-origin test under strict information cut-offs, and a small number of supporting decompositions, are specified as the remaining work in Section 7.4.

The remainder of the paper is organised in the conventional order. Section 2 establishes the background in the public record, which is the calibration target. Section 3 sets out the method in a single section. Section 4 gives the data and assumptions for both networks and the coupling between them. Section 5 presents the results across the six lenses. Section 6 reports validation, the structural-robustness and breach tests, and the internal audit. Section 7 discusses objections, transferability, recommendations and the full set of limitations. Section 8 concludes on the method. Four appendices provide a glossary, the full configuration of both networks, a note on distributional shape, and the component-breach detail.

An earlier paper by the present author, Dhand (2026), put the clause 4.2.3 evidence to this journal's readers and illustrated the dependence argument on a seven-node, author-estimated table using an earlier build of the same simulation engine. That treatment was illustrative, and it claimed neither calibration to the primary source nor validation. The present paper is a distinct study and differs in every respect that carries its claims. The registers are rebuilt to Phase 1 from Table 16 of the 2013 report rather than estimated. The schedule domain and its explicit coupling to cost are new. The rolling-origin forward test of Section 5.7, and the topology, distributional, component-breach and multi-seed tests of Section 6, are new. The shared premise and the shared tool lineage are declared in Appendix B.

2. Background: the public record as calibration target

Two features of HS2's history bear directly on the modelling question and form the target against which the reconstructed model is later checked. The first is the trajectory of the Phase 1 cost forecast. The second is the structure of the reference class for comparable rail projects, which indicates the direction in which a frozen, independence-based estimate is expected to drift.

2.1 The trajectory of the Phase 1 forecast

The published Phase 1 figure did not move in a single discontinuity. It climbed in steps, each documented and each defended at the time it was current. Table 1 records the trajectory from the position in the 2013 risk model to the 2026 reset, drawing only on Government, National Audit Office and Parliamentary sources. The figures span more than one price base and more than one scope convention, and they are presented as a sequence of contemporaneous official positions rather than as a single deflated series; the caption states the caveat.

Date	Source	Phase 1 position	Note
Oct 2013	Cost and Risk Status Report, Table 16	Construction base, P95 of order £21bn	2011 prices; threats modelled as independent (cl. 4.2.3)
2016	NAO, HC 235	£27.4bn	Reported Phase 1 estimate
2020	Full Business Case	£44.6bn	£40.3bn target plus £4.3bn contingency; ben–cost ratio of order 1.2
Nov 2023	DfT six-monthly report (HCWS34)	Main civils £21.8–£23.4bn	A single disaggregated system figure
Jul 2024	NAO, HC 128	£45–£57bn	DfT and HS2 Ltd ranges that did not reconcile
Feb 2025	Public Accounts Committee, HC 357	2019 budget no longer viable	Inflation contribution of order £10bn
May 2026	**Accounting Officer Assessment**	**£87.7–£102.7bn**	**Reset; opening 2036–2043; spend to Mar 2026 £44.2bn**

Table 1. The Phase 1 cost forecast in the public record, from the 2013 risk-model position to the 2026 reset. Figures are contemporaneous official positions and are not deflated to a common price base or harmonised for scope changes; they are shown to establish direction and magnitude, not a continuous real-terms series. Sources: Department for Transport and HS2 Ltd (2013); National Audit Office (2016, 2024); Department for Transport and HS2 Ltd (2020); Department for Transport (2023); Committee of Public Accounts (2025); HS2 Ltd and Department for Transport (2026).

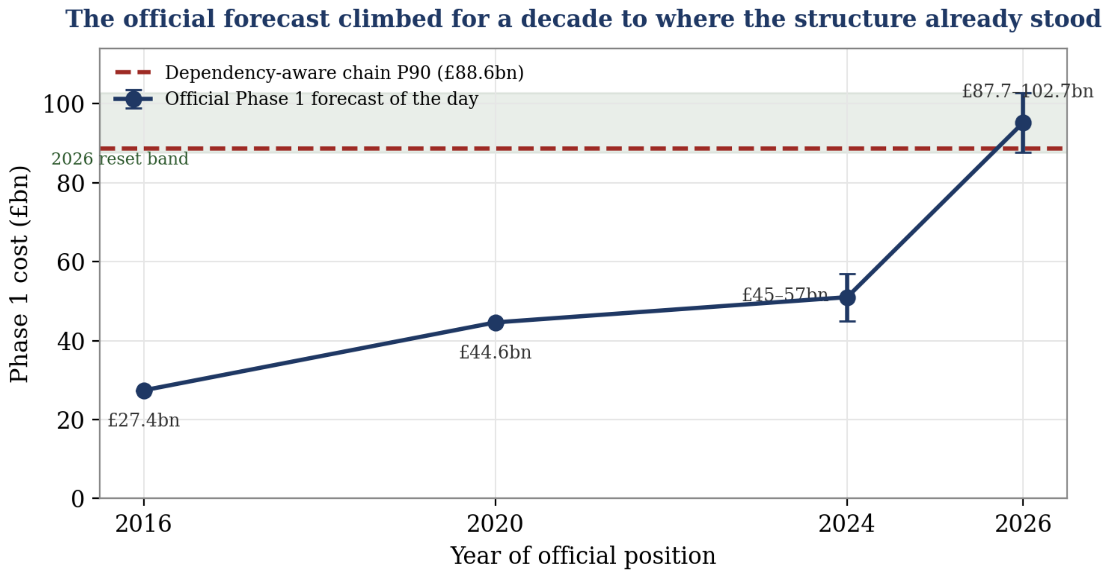


Figure 1. The Phase 1 cost forecast in the public record (Table 1), against the dependency-aware chain P90 of £88.6bn*. The official figure climbed from £27.4bn in 2016 to the £87.7–102.7bn reset of 2026, crossing the level the coupled register implied only at the very end of the decade. Where the contemporaneous position was a range, the marker is its midpoint and the bar its extent; the figures are contemporaneous official positions on differing price bases (see Table 1).*

The pattern in Table 1 is not a single shock but a sequence of upward revisions, each one defended until it was replaced by the next. This is the visible form of the frozen-baseline problem described in Section 1.2: not a sudden discovery but a steady accumulation, disclosed in instalments. Two narrower observations from the record matter for the cost model later in the paper. The November 2023 disaggregation of the main-civils package to between £21.8bn and £23.4bn (Department for Transport, 2023) provides a directly sourced magnitude for the single largest construction system, and the May 2026 reset provides both the cost band and the schedule slip from which the schedule-to-cost coupling is calibrated.

2.2 The reference class and the direction of drift

Whether a frozen, independence-based estimate is expected to drift upward or downward is not a matter of opinion; it can be read from the distribution of outcomes on comparable projects. For a large international sample of rail schemes, Flyvbjerg, Holm and Buhl (2002) report a mean cost overrun against the figure used at the decision to build of about 44.7 per cent, with a standard deviation of about 38.4 per cent. Two consequences follow arithmetically. First, a contingency of around a third (of the order HS2 carried, and of the order many programmes treat as a P90 allowance) sits not near the ninetieth percentile of the rail reference class but closer to its centre, around the fortieth to fiftieth percentile. Second, a genuine P90 in that reference class requires an uplift of approximately $44.7 + 1.28 \times 38.4 \approx 94$ per cent, not a third. A contingency labelled as a high-confidence allowance was, against the relevant comparison set, close to a central estimate.

This is the planning-fallacy signature: estimates anchored to the specifics of the project at hand, rather than to the distribution of outcomes for its reference class, are systematically optimistic (Kahneman, 2011). The reference-class parameters used above are two decades old, and a more recent synthesis reports a smaller central figure for some categories (Flyvbjerg and Gardner, 2023); the point relied upon here is only the direction and rough scale of the drift, which both sources agree is upward and large for rail. The reference class identifies the likely destination of the estimate. It does not, on its own, identify the route by which a particular project will reach it. That is the task of the network model in Sections 3 to 5: not to assert a final figure, but to let the project's own risks influence one another and to read off where the resulting distribution falls.

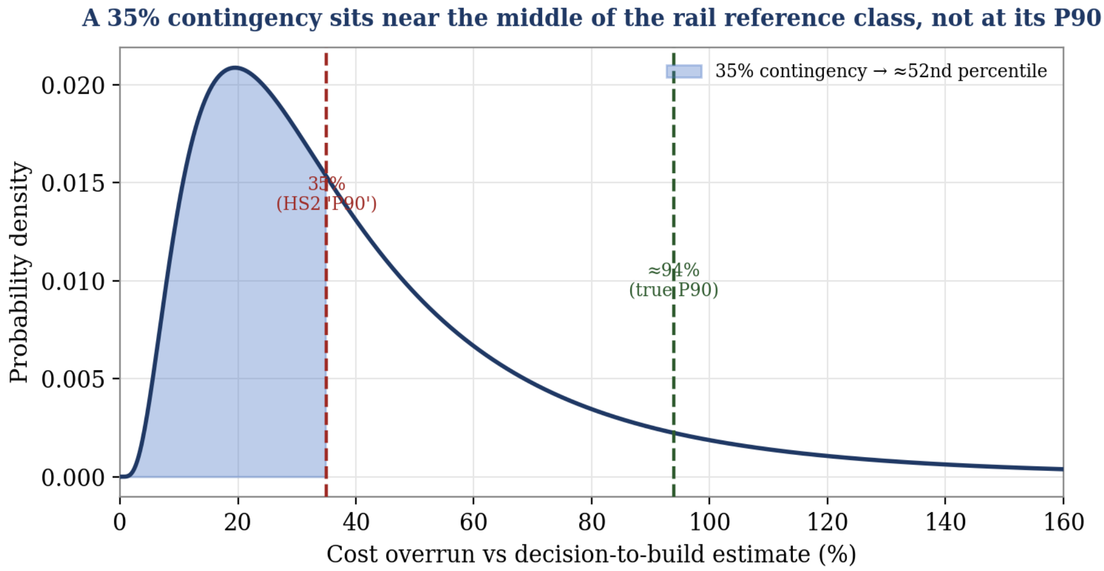


***Figure 2. The rail reference class of Flyvbjerg, Holm and Buhl (2002)**: a lognormal fitted to a mean overrun of 44.7 per cent with a standard deviation of 38.4 per cent. A 35 per cent contingency (red) falls near the 52nd percentile, not the 90th; a genuine P90 in this class requires an uplift of approximately 94 per cent (green). The shaded area is the probability mass below 35 per cent.*

One element of the background connects directly to a later test. A natural question is whether the historic Phase 1 forecasts also breached, system by system, the ranges a dependency-aware model would assign; a component-level test of the model against the record. That test requires the per-system, per-iteration output of the simulation. The per-node output has now been exported for the independent and the modelled runs, and the test is performed in Section 6.8 and Appendix D; the breach pattern observed in Table 2 is the qualitative counterpart it makes precise.

3. Method

The method is presented in one section, as a single connected procedure. It has five parts: the representation of risk as a network (3.1); the rule by which several causes combine on a single effect (3.2); the handling of statistical dependence and the deliberate separation of two distinct channels through which risks can move together (3.3); the mechanism that couples the schedule network to the cost network (3.4); and the risk measures and simulation procedure, including the four configurations that isolate the independence effect (3.5).

3.1 Risk as a Bayesian network

A Bayesian network represents a set of uncertain quantities as nodes in a directed acyclic graph, with an arrow drawn from one node to another where the first bears causally on the second (Pearl, 1988; Jensen and Nielsen, 2007). Each node carries a marginal probability that its risk is realised and, conditional on realisation, a magnitude. Each arrow carries a conditional dependence: the probability that the child risk fires is higher when its parent has fired than when it has not. The structure encodes the engineering logic of the programme directly (that ground conditions bear on tunnelling, that tunnelling bears on the station programme, that funding instability bears on contractor performance) rather than leaving such relationships to be approximated by a correlation coefficient applied after the fact. The application of this apparatus to project and cost risk follows Fenton and Neil (2018).

Two networks are built. The schedule network has ten systems and thirteen arrows; its node magnitudes are measured in days of critical-path slip, and its output is the total programme slip. The cost network has twelve systems and fifteen arrows; its node magnitudes are measured in pounds, and its output is the total Phase 1 cost. The two are not run in isolation. One node of the cost network, Schedule Delay Cost, takes its magnitude from the output of the schedule network, by the mechanism set out in 3.4. The full node and arrow specifications for both networks are in Appendix C.

3.2 Combining several causes: the noisy-OR rule

Where a node has several parents, the model requires a rule for combining their influence. A naïve rule that simply added or multiplied the parental effects would let a node with many parents reach implausible certainty. The model instead uses the noisy-OR formulation (Fenton and Neil, 2018), in which each parent that fires makes an independent attempt to trigger the child, and the child fails to fire only if every firing parent independently fails to trigger it. If parent i contributes an activation probability p_i, then for the set F of parents that have fired,

$$P(\text{child fires} \mid \text{parents } F) = 1 - \prod_{i \in F} (1 - p_i) \quad ,$$

This has the properties an engineer expects. A single firing parent raises the child's probability by a bounded amount; additional firing parents raise it further but with diminishing return; and the probability approaches, but never exceeds, certainty. The effect is that a node sitting downstream of several active hazards becomes very likely, though not guaranteed, to fire; which is the realistic behaviour of, for example, a main-civils package exposed simultaneously to adverse ground, utilities, supply-chain and scope pressure. The noisy-OR rule is enabled in all networked runs reported here.

3.3 Two channels of dependence, and why they are kept separate

Risks can move together for two quite different reasons, and the method keeps them strictly apart because conflating them is a common source of double counting. The first channel is structural: one risk raises the probability of another through a documented causal arrow, as in 3.1. The second is residual statistical correlation: two risks tend to co-move for reasons not captured by any arrow (a shared exposure to weather, to a labour market, to a macroeconomic cycle) and this residual is handled, in the standard way, by inducing rank correlation between the sampled inputs using the Iman and Conover (1982) method.

These are distinct mechanisms and must not both be asked to carry the same relationship. In this study the structural channel does the work, by design. The residual rank-correlation stage carries no designed coefficients between systems; the Iman–Conover machinery is run so that the model is demonstrably capable of imposing residual correlation, but it is not used to encode the dependence, which is carried entirely by the documented arrows. Section 6.4 confirms the consequence numerically: switching the correlation stage off versus on moves the cost P90 by only about £0.003bn, and the schedule P90 by about six days; under one per cent of the independence gap in each domain. Every material pound and day of the dependence effect in Section 5 therefore arises from the causal structure, not from a tuned correlation. The known hazard that linear correlation behaves poorly in the tails of skewed or heavy-tailed inputs (Embrechts, McNeil and Straumann, 2002) is noted because it would constrain any future use of the residual channel, not because it is engaged here.

3.4 Coupling the schedule network to the cost network

The two networks are joined at a single, fully documented point, so that schedule risk enters the cost answer through a mechanism that can be inspected and disputed rather than through an assumed correlation. The schedule network produces a distribution of programme slip in days. That slip is converted into a cost by an extended-overhead, or standing-army, argument: a programme that runs longer continues to incur its time-related overheads (client and management organisation, site establishment, financing of works in progress) for the additional duration. The converted figure becomes the magnitude of the cost node HC12, Schedule Delay Cost, which then propagates within the cost network to inflation and funding (HC08) and beyond.

The derivation is built from sourced quantities wherever the record supplies them, and the single modelled fraction is stated as such. Reported spend to March 2026 is £44.2bn (HS2 Ltd and Department for Transport, 2026); spread across approximately six years of active construction since notice to proceed, this is a gross burn of about £7.37bn per year. Not all of that burn is duration-sensitive; the share that is (the time-related overhead rather than the fixed quantities of concrete and steel) is taken at 13.6 per cent, within the 10–20 per cent range typical of construction economics for programmes of this kind, giving a central duration-sensitive rate of about £1.0bn per year. The slip itself is taken from the reset: a central slip of thirteen years (the midpoint of the official ten-to-seventeen-year band) and an upper slip of seventeen years (its top). The resulting node ranges from about £2.0bn, through a most-likely £13.0bn, to about £25.5bn, the upper tail using an escalated rate to reflect overhead inflation across a very long extension. The full arithmetic is tabulated in Section 4.3 and Appendix C. The 13.6 per cent overhead share is the single load-bearing modelling assumption in the coupling, and it is flagged as such in Section 7.4; a floor check against the 2013 client-and-corporate overhead line, which implies an order-of-magnitude lower bound near £0.2bn per year, confirms the rate is not implausibly high.

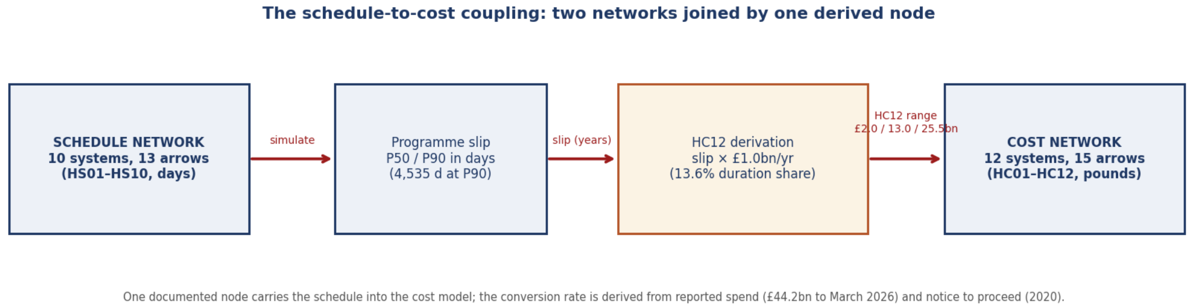


***Figure 2a. The schedule-to-cost coupling drawn as a pipeline**: the schedule network's simulated slip enters the cost network through the single derived node HC12, converted at the extended-overhead rate of about £1.0bn per duration-sensitive year. The full derivation, with each line marked sourced, derived or modelled, is tabulated in Section 4.3.*

3.5 Risk measures, simulation, and the four configurations

Each network is evaluated by Monte Carlo simulation. Per-node magnitudes are drawn from PERT (beta-PERT) distributions specified by a minimum, most-likely and maximum value, with the standard shape parameter $\lambda = 4$, sampled by Latin-hypercube to improve coverage of the input space (Vose, 2008). Every run reported in the body uses a fixed random seed of 4209 and 100,000 iterations, so that differences between runs are differences of model configuration and not of sampling noise. From each output distribution the analysis reads the median (P50), the P80 and P90 percentiles, and, for the tail, the conditional value at risk (the mean of outcomes beyond the P90, denoted CVaR(90%)) which captures the severity of the tail rather than only its threshold (Rockafellar and Uryasev, 2000). Comparisons between configurations are made at matching percentiles, which is the appropriate basis when two distributions differ in shape as well as location (Müller and Stoyan, 2002).

To isolate the independence effect cleanly, each network is run in four configurations. The differences between them are differences of permitted dependence only; the nodes, ranges, seed and iteration count are identical throughout.

The first configuration is the **independent** register: both dependence channels off, no arrows and no residual correlation. This is the 2013-style model, and it is the baseline against which the independence error is measured. The second is **correlation only**: the independent register with the residual rank-correlation stage switched on and the arrows still off. Compared against the same register with the stage off, this isolates the contribution of the correlation channel, which Section 6.4 shows to be negligible, about £0.003bn on cost, so that the dependence effect cannot be attributed to it. The third is the **dependency-aware chain**: the documented arrows switched on under the noisy-OR rule, with the correlation stage present but still targeting zero. This is the model that lets the project's risks influence one another. The fourth is the **updated** chain: node probabilities and ranges revised on evidence, first on the public record (the *official* lens) and then on the system-level magnitudes a sponsor already holds (the *modelled* lens). This is the baseline allowed to learn.

Applied to the schedule network and the cost network in turn, these configurations generate the six lenses of Section 5: the independent and dependency-aware schedule; and the independent, dependency-aware, official and modelled cost. The independence error in each domain is the difference between the dependency-aware P90 and the independent P90, with all magnitudes held fixed. The dependence effect each lens produces can itself be separated into two mechanisms (the conditional structure that the arrows impose, and the noisy-OR rule that combines several active parents at a shared child) by re-running the chain with the noisy-OR rule disabled. Section 6.5 reports that decomposition for both networks.

3.6 Application to any project

The method transfers without the engine used here; a spreadsheet, a reference class and a willingness to draw arrows suffice. The register is first collapsed to the eight to fifteen systems that describe how the asset is actually built, because three hundred rows conceal the structure that a dozen systems reveal. For every pair of systems one question is then asked: if the first fires, does the second become more likely? Where the answer is yes an arrow is drawn, and only the genuine relationships are retained, so that the result is a single page a review meeting can interrogate. Each system is given a probability in isolation and a three-point magnitude anchored to a reference class, so that optimism is bounded; where an in-house figure disagrees with the class, the burden of justification falls on the project, not on the auditor. Each arrow is then given its lift, the probability of the child given that the parent has fired. That single value is the dependency an independent model discards; it is the load-bearing input, and it must be defensible one arrow at a time.

The register is then run twice, once with the arrows inactive and once with them active, and the difference between the two P90 values is the project's independence error, in pounds or in days. The variance decomposition of the dependency-aware run locates the most connected joint, which is where a mitigation pound earns most (Section 5.5). Finally, the intelligence log is maintained from the first day: each material disclosure updates the relevant prior or magnitude, the model is re-run, and the movement is recorded with its source, so that the baseline drifts toward the realised position as evidence arrives and the closing position of one project becomes the opening prior of the next. Sections 5.6 and 5.7 demonstrate this procedure end to end on HS2's own registers.

4. Data and assumptions

This section specifies the two registers, the coupling between them, and the evidence used to update each network. The complete configuration (every node and every arrow for both networks) is reproduced in Appendix C; the tables here give the node magnitudes, which carry the substance of the model, and summarise the dependency structure. Throughout, a distinction is maintained between quantities taken from the public record and quantities that are modelled, and the modelled quantities are identified as such.

4.1 The schedule register

The schedule network has ten systems, listed in Table 2 with their marginal probability of firing and their three-point range in days of critical-path slip. The probabilities and the relative ordering of the systems are anchored to the programme's own risk documentation and to the National Audit Office's progress reporting (HS2 Ltd, 2012, Appendices B and D; National Audit Office, 2020). The day-magnitudes themselves, however, are **modelled**: the public record does not attribute the programme slip to individual systems, and only the total slip is independently sourced, from the 2026 reset. This is the schedule analogue of the disclosure gap discussed for cost in Section 7, and it is recorded plainly here; the shape and ordering of the schedule register are evidence-based; the per-system day counts are an informed allocation, and the result that survives this caveat is the structural gap (Section 5.1), which depends on the ordering and not on the absolute calibration.

ID	System (schedule)	P(fire)	Min	M.L.	Max
HS01	Land and property acquisition	0.70	90	365	900
HS02	Ground conditions	0.65	90	270	730
HS03	Utilities diversions	0.50	90	270	640
HS04	Main civils construction (critical path)	0.65	180	540	1460
HS05	Stations (Euston redesign / pause)	0.50	180	540	1825
HS06	Environmental clearance and consents	0.55	120	365	900
HS07	Contractor performance	0.45	120	365	1095
HS08	Scope and design changes	0.45	90	270	730
HS09	Governance, plan revision and reset	0.50	90	365	1095
HS10	Funding instability (spending-review profile)	0.55	180	540	1460

***Table 2. The ten-system schedule register**. Magnitudes are in days of critical-path slip. Probabilities and system ordering are anchored to HS2 Ltd (2012) and National Audit Office (2020); per-system day-ranges are modelled (the public record attributes slip only in total). Thirteen dependency arrows connect these systems; they are listed in Appendix C.2.*

4.2 The cost register

The cost network has twelve systems, listed in Table 3 with probabilities and three-point ranges in pounds. Eleven of the twelve are conventional cost systems; the twelfth, HC12 (Schedule Delay Cost), is the coupling node and is derived in 4.3 rather than assigned directly. The magnitudes here are better anchored than the schedule's, because the cost record discloses more. The base figures derive from the programme's 2012 and 2013 cost documentation (HS2 Ltd, 2012; Department for Transport and HS2 Ltd, 2013, Table 16), and the largest system, main civils, has its central and upper range informed by the November 2023 disaggregation to between £21.8bn and £23.4bn (Department for Transport, 2023). Three systems (utilities (HC03), stations (HC05) and railway systems (HC06)) have no separately published Phase 1 magnitude and are therefore **not calibrated** to a public figure; their ranges are conservative placeholders, which makes the modelled cost answer of Section 5 an understatement rather than an overstatement on those three systems. This is flagged in Section 7.4.

ID	System (cost)	P(fire)	Min	M.L.	Max
HC01	Land and property	0.80	1.6	3.0	6.0
HC02	Ground conditions	0.65	0.5	2.0	6.0
HC03	Utilities and enabling works †	0.50	0.5	1.5	4.0
HC04	Main civils (tunnels, earthworks, structures)	0.65	6.3	18.0	30.0
HC05	Stations and depots †	0.45	3.3	8.0	16.0
HC06	Railway systems (track, OHLE, signaling) †	0.40	2.0	4.5	9.0
HC07	Environmental and consents	0.70	0.3	2.0	6.0
HC08	Inflation and funding	0.50	1.0	10.0	24.0
HC09	Contractor and supply chain	0.45	0.8	7.0	16.0
HC10	Scope and design	0.45	0.6	7.0	17.0
HC11	Governance and client	0.50	2.4	5.0	10.0
HC12	**Schedule delay cost (derived; see 4.3)**	**0.55**	**2.0**	**13.0**	**25.5**

***Table 3. The twelve-system cost register**. Magnitudes are in £bn. † marks the three systems with no separately published Phase 1 figure, whose ranges are conservative and uncalibrated. HC12 is derived in Section 4.3. Fifteen dependency arrows connect these systems; they are listed in Appendix C.4. Sources: HS2 Ltd (2012); Department for Transport and HS2 Ltd (2013); Department for Transport (2023).*

4.3 The schedule-to-cost coupling (node HC12)

The coupling node converts the schedule network's programme slip into a cost, by the extended-overhead argument of Section 3.4. Table 4 sets out the full derivation in three steps, marking each quantity as sourced, derived or modelled. Only one quantity is modelled: the share of the gross programme burn that is duration-sensitive, taken at 13.6 per cent. The annual rate, the slip band, and the spend are otherwise sourced or follow arithmetically.

Step / item	Value	Basis and status
Step 1; cost of one year of delay		
HS2 spend to end March 2026	£44.2bn	AO Assessment May 2026; sourced
Construction start (notice to proceed)	2020	AO Assessment May 2026; sourced
Elapsed construction years to cut-off	6 yr	Derived (2026 − 2020)
Average gross programme burn	£7.37bn/yr	Derived (spend / years)
Extended-overhead share of gross burn	*13.6%*	*Modelling assumption (10–20% range); modelled*
Central duration-sensitive rate (used)	£1.00bn/yr	Derived (burn × share)
Cross-check: client-side overhead floor	£0.19bn/yr	2013 Table 16 (contracts & delivery overhead); sourced floor
Step 2; programme slip		
Slip (low (optimistic prior)	2 yr	Optimistic prior tail) modelled
Slip (central	13 yr	AO Assessment midpoint (2036–2043)) sourced
Slip (high	17 yr	AO Assessment upper bound (2043)) sourced
Step 3; HC12 = slip × rate		
Minimum	£2.0bn	2 yr × central rate
Most likely	£13.0bn	13 yr × central rate
Maximum	£25.5bn	17 yr × escalated tail rate (£1.50bn/yr)

Table 4. Derivation of the coupling node HC12, Schedule Delay Cost, from programme slip. *The single modelled quantity (the 13.6% duration-sensitive overhead share) is shown in italics; it is the load-bearing assumption of the coupling and is flagged in Section 7.4. The escalated tail rate in the maximum reflects overhead inflation across a very long extension. Sources: HS2 Ltd and Department for Transport (2026); Department for Transport and HS2 Ltd (2013).*

Two checks bound this derivation. The floor check compares the modelled central rate of about £1.0bn per year against the 2013 client-and-corporate overhead line, which implies a duration-sensitive lower bound near £0.19bn per year; the modelled rate sits an order of magnitude above the bare client-side floor, as expected once site establishment and works financing are included, and well below the gross burn. The upper check is the reset itself: a seventeen-year slip at the escalated rate produces about £25.5bn, which is the order of magnitude by which the cost band moved between the 2020 business case and the 2026 reset. The node is therefore neither trivial nor inflated; it is commensurate with the disclosed movement.

4.4 Configurations and the evidence for updating

Each network is run in the four configurations of Section 3.5. The independent, correlation-only and dependency-aware runs share the registers of Tables 2 and 3 unchanged; they differ only in permitted dependence. The two updated runs (the official and modelled lenses) revise the node probabilities and observed magnitudes against evidence, and the evidence is documented system by system in observation registers released alongside this paper.

The *official* lens admits only what the public record states. For each system it records whether the record shows the risk as having materialised and, where a number is published, the observed magnitude; many systems are recorded as fired with no separately published figure, sourced to the National Audit Office's reporting (National Audit Office, 2020), the 2013 status report, and the Department's six-monthly reports, including the July 2025 disclosure of a land liability of order £0.6bn (Department for Transport, 2025). The *modelled* lens admits, in addition, the system-level magnitudes a sponsor already holds (for example a land-and-property impact of order £5bn and a ground-conditions impact of order £3.5bn, each anchored to the 2013 base and to later disclosures) while leaving the three uncalibrated systems (HC03, HC05, HC06) at their conservative placeholders. The official lens is the public record allowed to speak; the modelled lens is the sponsor's own knowledge allowed to speak. Neither introduces information unavailable at the time; both simply allow the baseline to learn.

5. Results: the six lenses

The results are presented as six lenses on the same programme. The first two view the schedule: independent, then dependency-aware. The next four view the cost: independent, dependency-aware, then updated on the public record (official) and on the sponsor's own magnitudes (modelled). Between the schedule and cost lenses sits the coupling of Section 4.3, by which the schedule's slip becomes a cost. The independence error in each domain is read as the difference between the dependency-aware P90 and the independent P90, with all magnitudes held fixed.

5.1 Lenses one and two: the schedule, independent and dependency-aware

The schedule network is shown in Figure 3: ten systems, thirteen arrows, with funding instability and governance at the head of the structure and contractor performance at its foot, where five arrows converge. The convergence is not incidental; it is the engineering reality that contractor performance is degraded by funding instability, by land and consents delay, by utilities, by scope churn and by an unstable design simultaneously.

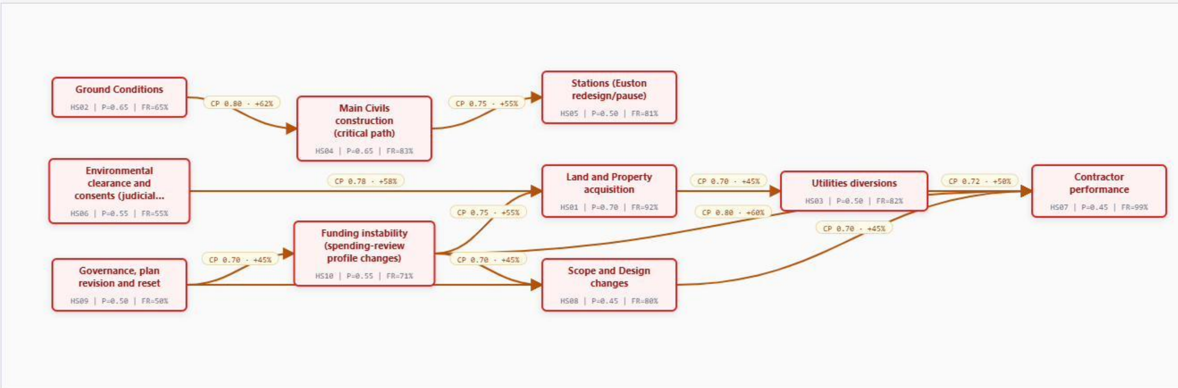


Figure 3. The ten-system schedule network. *Arrows run from a system to those whose probability of slipping it raises. Contractor performance (HS07) is the convergence node, with five parents; funding instability (HS10) and governance (HS09) are sources.*

Running the network independently (every arrow off) reproduces the behaviour of a 2013-style schedule risk model and gives a P90 programme slip of 3,671 days, about 10.1 years. Switching the thirteen documented arrows on, with the noisy-OR rule, and changing nothing else (not a probability, not a day-range, not the seed) raises the P90 to 4,535 days, about 12.4 years. The difference, **864 days, or roughly 2.4 years**, is schedule exposure attributable to model structure alone. Figure 4 overlays the two distributions; the dependency-aware curve is shifted bodily to the right and is also less skewed, the right tail having been pulled in as downstream systems become reliably, rather than occasionally, active.

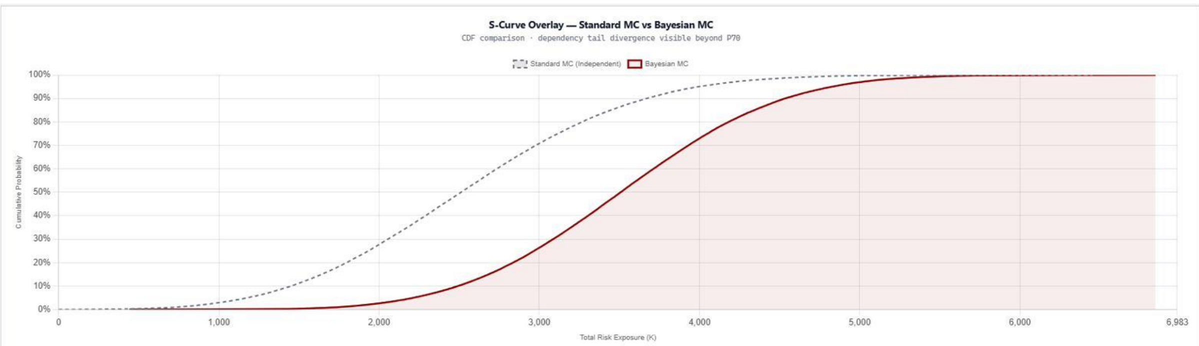


Figure 4. Cumulative distribution of total programme slip (days), independent versus dependency-aware. *Switching on the documented dependencies shifts the whole curve to the right; the P90 moves from 3,671 to 4,535 days.*

Table 5 sets out the schedule percentiles for all four configurations. The dependency-aware chain (12.4 years at P90) falls inside the official reset of a 2036–2043 opening; a slip of ten to seventeen years (HS2 Ltd and Department for Transport, 2026). The updated lenses are the 2026 endpoints of the rolling-origin runs of Section 5.7, in which each document is admitted only at the cut-off on or after its date: on the public record alone the P90 reaches 4,671 days (12.8 years), and on the sponsor-grade magnitudes

6,187 days (16.9 years). All three dependency-aware lenses therefore stand inside the ten-to-seventeen-year band the reset later confirmed, and they bracket the coupling of Section 4.3 from both sides: the central thirteen-year slip used for HC12 sits between the chain (12.4) and the official lens (12.8), and the seventeen-year top from which the coupling's upper bound draws is the region the modelled lens reaches.

Schedule lens	P50	P80	P90	CVaR90	P90 (years)
Independent (2013-style)	2,512	3,267	3,671	4,097	10.1
Dependency-aware (chain)	**3,507**	**4,182**	**4,535**	**4,908**	**12.4**
Updated 2026; official (forward log)	3,670	4,325	4,671	–	12.8
Updated 2026; modelled (forward log)	**5,002**	**5,805**	**6,187**	**–**	**16.9**

Table 5. Programme slip across the four schedule configurations, in days, with the P90 also in years. *The independence error is the gap between the first two rows: 864 days, about 2.4 years (bootstrap standard error about 7 days). The updated rows are the 2026 endpoints of the released rolling-origin logs, which publish P50, P80 and P90; the tail average for these lenses is reported from the five-seed study in Section 6.10. The chain, official and modelled lenses (12.4, 12.8 and 16.9 years) all stand inside the official reset band of ten to seventeen years.*

The mechanism behind the shift is visible in the firing behaviour of the nodes. Table 6 reports, for each system, its standalone probability and the rate at which it actually fires once the network is active. The downstream systems do not merely become a little more likely; they become near-certain. Contractor performance rises from a base of 0.45 to 0.986, and main civils, land, utilities, stations and scope all settle well above their standalone probabilities, while the three source systems (ground, environmental, governance) fire at their base rates, as they must, having no parents. This is the chain doing exactly what an independent model forbids: letting an active hazard raise the odds of the next.

ID	System	Base P	Fires at	Parents
HS07	Contractor performance	0.45	0.99	HS10, HS01, HS06, HS03, HS08
HS01	Land and property	0.70	0.92	HS10, HS06
HS04	Main civils	0.65	0.83	HS02
HS03	Utilities diversions	0.50	0.82	HS01
HS05	Stations	0.50	0.81	HS04
HS08	Scope and design	0.45	0.80	HS10, HS09
HS10	Funding instability	0.55	0.71	HS09
HS02	Ground conditions	0.65	0.65	(source node)
HS06	Environmental and consents	0.55	0.55	(source node)
HS09	Governance and reset	0.50	0.50	(source node)

Table 6. Standalone probability versus realised firing rate for each schedule system in the dependency-aware run. *Systems are ordered by how far they move. The three source systems fire at their base rate; downstream systems become near-certain. Rates are from the engine's node diagnostics and are consistent with the recomputed mean (Section 6).*

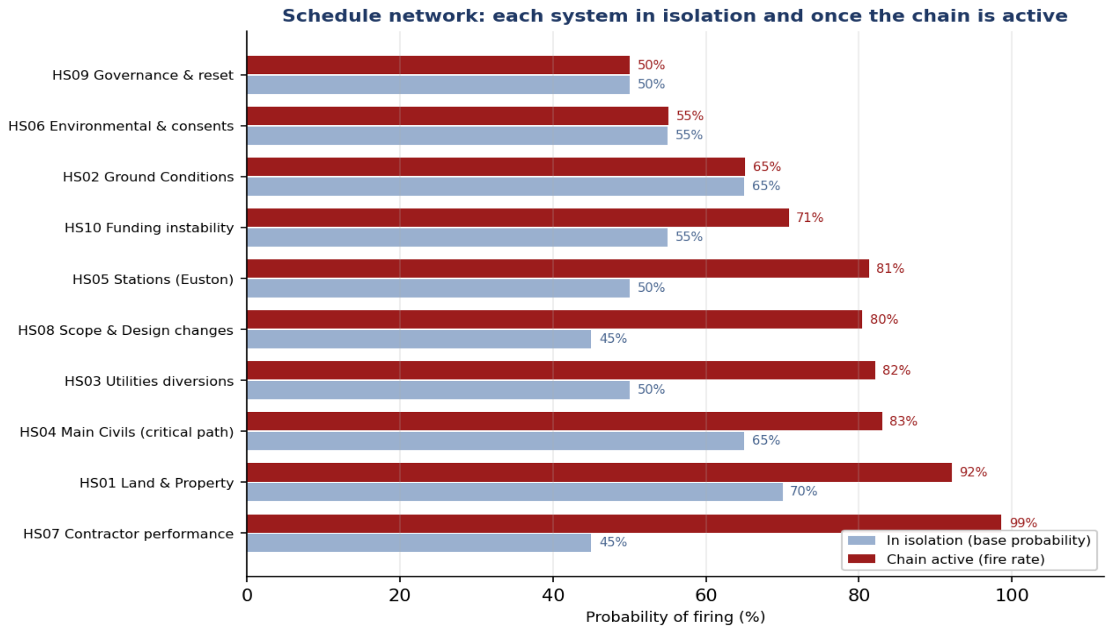
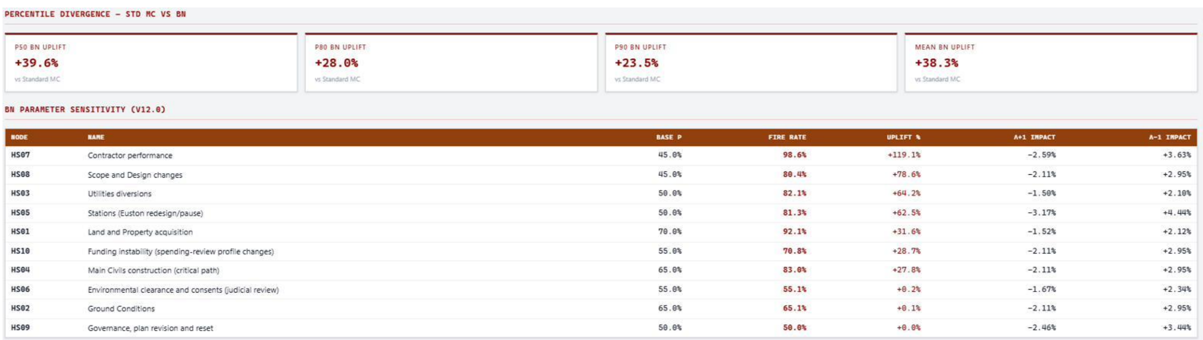


Figure 4a. The schedule mechanism drawn*: each system's standalone probability against its realised firing rate once the chain is active, from the engine's node diagnostics. The three source systems are unchanged; contractor performance, the convergence node, rises from 0.45 to 0.986.*


PERCENTILE DIVERGENCE – STD MC VS BN

P50 BN UPLIFT	P80 BN UPLIFT	P90 BN UPLIFT	MEAN BN UPLIFT
+39.6%	+28.0%	+23.5%	+38.3%
vs Standard MC	vs Standard MC	vs Standard MC	vs Standard MC

BN PARAMETER SENSITIVITY (V12.0)

NODE	NAME	BASE P	FIRE RATE	UPLIFT %	A+1 IMPACT	A-1 IMPACT
HS07	Contractor performance	45.0%	98.6%	+119.1%	-2.59%	+3.63%
HS08	Scope and Design changes	45.0%	80.4%	+78.6%	-2.11%	+2.95%
HS03	Utilities diversions	50.0%	82.1%	+64.2%	-1.50%	+2.10%
HS05	Stations (Euston redesign/pause)	50.0%	81.3%	+62.5%	-3.17%	+4.44%
HS01	Land and Property acquisition	70.0%	92.1%	+31.6%	-1.52%	+2.12%
HS10	Funding instability (spending-review profile changes)	55.0%	70.8%	+28.7%	-2.11%	+2.95%
HS04	Main Civils construction (critical path)	65.0%	83.0%	+27.8%	-2.11%	+2.95%
HS06	Environmental clearance and consents (judicial review)	55.0%	55.1%	+0.2%	-1.67%	+2.34%
HS02	Ground Conditions	65.0%	65.1%	+0.1%	-2.11%	+2.95%
HS09	Governance, plan revision and reset	50.0%	50.0%	+0.0%	-2.46%	+3.44%

Figure 4b. Percentile divergence on the schedule network, from the engine's comparison dashboard: *the dependency-aware run against the independent run at matching percentiles. The P50 rises by 39.6 per cent, the P90 by 23.5 per cent and the mean by 38.3 per cent; the uplift is largest in the body of the distribution, which is the relocation signature discussed in Section 5.3.*

The dependence also rearranges which systems drive the uncertainty, though less dramatically than on the cost side. Under independence the three largest variance contributors to the slip are stations, funding instability and main civils. Under the chain, stations remains first and funding second, but contractor performance falls sharply as a variance driver (from about 8 per cent to about 5 per cent) precisely because it has become near-certain to fire and therefore stops varying. A system that always happens cannot be a source of variability; it becomes part of the baseline. This is a recurring and slightly counter-intuitive feature of dependency-aware models, and it appears in stronger form on the cost network in Section 5.3.

5.2 The bridge realised: schedule becomes cost

Before turning to the cost lenses, it is worth stating what the schedule result implies for cost, because the two are joined. The dependency-aware schedule places the central programme slip in the region of twelve to thirteen years. Through the coupling of Section 4.3, at a duration-sensitive rate of about £1.0bn per year, a slip of that order carries an extended-overhead cost with a most-likely value near £13bn; which is exactly the magnitude assigned to node HC12. The schedule network is not a separate study that happens to sit alongside the cost network; it is the origin of one of the cost network's twelve systems, and a material one. In the dependency-aware cost run, HC12 fires at 0.98 and carries about 17 per cent of the cost variance (Section 5.3). The schedule's slip is therefore not an external narrative imposed on the cost model; it enters the cost answer as a quantified, sourced node, and it does so as one of the cost model's larger contributors.

5.3 Lenses three and four: the cost, independent and dependency-aware

The cost network is shown in Figure 5: twelve systems, fifteen arrows. Its shape differs from the schedule's in a way that matters. Main civils (HC04) is a convergence node fed by four parents (ground, utilities, contractor and scope) and it in turn feeds stations and the coupling node HC12; HC12 then feeds inflation, which feeds railway systems. The structure carries a long propagation path from the early-works systems, through main civils and the schedule-delay cost, into inflation, which is the cost network's analogue of the schedule's contractor convergence.

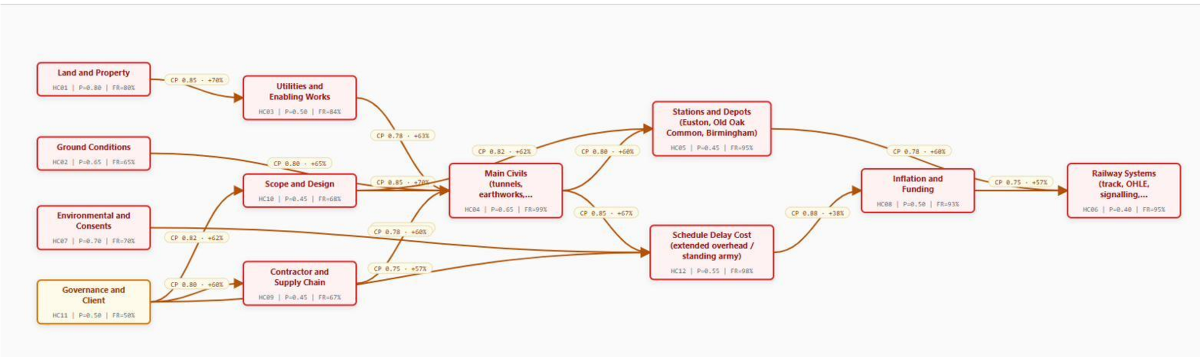


Figure 5. The twelve-system cost network. *Main civils (HC04) is the principal convergence node, with four parents; the coupling node HC12 (Schedule Delay Cost) sits downstream of*

main civils and feeds inflation (HC08). Governance and client (HC11) is a source feeding scope, contractor and the coupling node.

Run independently, the cost network gives a P90 of £67.4bn. Switching on the fifteen documented arrows, with the noisy-OR rule and the magnitudes held exactly as in Table 3, raises the P90 to £88.6bn. The difference, **approximately £21.3bn**, is cost exposure attributable to model structure alone, before any new information about the project is introduced. A bootstrap of the raw output places the sampling error on that figure at about £0.1bn (Section 6), so it is a structural shift, not a sampling artefact. Figure 6 overlays the two cost distributions; as on the schedule side, the dependency-aware curve moves bodily to the right and its right tail is drawn in, the median rising more than the P90 as the body of the distribution catches up with what was previously a thin tail.

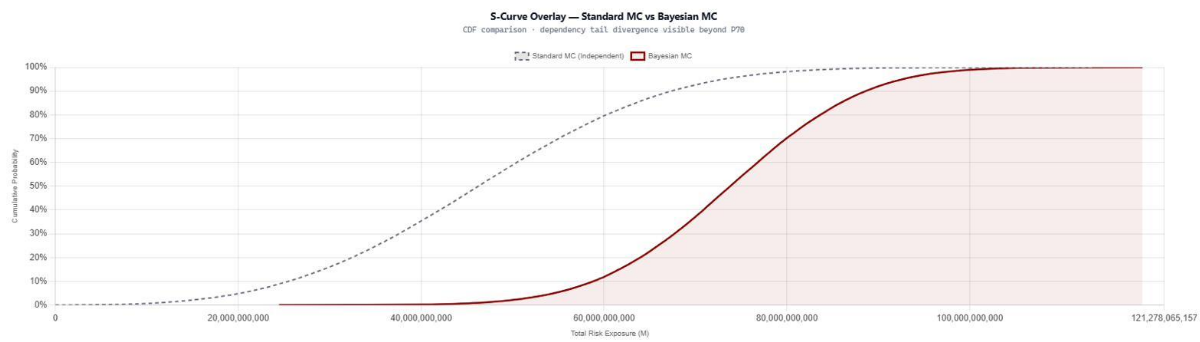


Figure 6. Cumulative distribution of total Phase 1 cost (£), independent versus dependency-aware. *The P90 moves from £67.4bn to £88.6bn; the median moves further still, from £46.4bn to £73.8bn.*

Table 7 sets the two cost distributions side by side. The median shift (£27.6bn) exceeds the P90 shift (£21.3bn), and the distribution changes shape as well as position: under independence it is mildly right-skewed, with a P90-to-P50 ratio of 1.46; under the chain it is very slightly left-skewed, with that ratio compressed to 1.20. The independent model expressed its risk as a long thin tail above a comfortable median; the dependency-aware model relocates that risk into the body of the distribution, where it is far harder to discount as a remote contingency. This is the practical danger of the independence assumption: it does not merely lower the headline figure, it reassures, by presenting accumulated risk as if it were a rare excursion.

Cost measure (£bn)	Independent	Dependency-aware	Shift
Median (P50)	46.25	73.81	+27.56
P80	60.24	83.66	+23.42
P90	**67.38**	**88.63**	**+21.25**
P95	73.04	92.69	+19.65
Tail mean, CVaR(90%)	74.81	93.89	+19.08
Mean	46.42	73.74	+27.32

Cost measure (£bn)	Independent	Dependency-aware	Shift
Skewness	+0.09	−0.05	falls to zero
P90 / P50 ratio	1.46	1.20	compresses

Table 7. The cost distribution under independence and under the dependency-aware chain, magnitudes held fixed (both runs at 100,000 iterations). *The independence error at P90 is £21.25bn (±£0.10bn bootstrap). The fall in skewness and the compression of the P90/P50 ratio show risk moving from the tail into the body of the distribution.*

The firing behaviour again drives the result. Table 8 gives each cost system's standalone probability and its realised firing rate in the chain. Main civils rises from 0.65 to 0.987; the coupling node HC12 rises from 0.55 to 0.979; stations and railway systems both settle above 0.94; inflation rises to 0.93, pulled up by the schedule-delay cost feeding it. The four source systems (land, ground, environmental, governance) fire at their base rates. The cost network is more strongly coupled than the schedule network: more of its nodes have multiple parents, and the convergence on main civils and on inflation is correspondingly sharper.

ID	System	Base P	Fires at	Parents
HC04	Main civils	0.65	0.99	HC02, HC03, HC09, HC10
HC12	Schedule delay cost	0.55	0.98	HC04, HC07, HC11
HC06	Railway systems	0.40	0.95	HC05, HC08
HC05	Stations and depots	0.45	0.95	HC04, HC10
HC08	Inflation and funding	0.50	0.93	HC12
HC03	Utilities and enabling	0.50	0.84	HC01
HC01	Land and property	0.80	0.80	(source node)
HC07	Environmental and consents	0.70	0.70	(source node)
HC10	Scope and design	0.45	0.68	HC11
HC09	Contractor and supply chain	0.45	0.67	HC11
HC02	Ground conditions	0.65	0.65	(source node)
HC11	Governance and client	0.50	0.50	(source node)

Table 8. Standalone probability versus realised firing rate for each cost system in the dependency-aware run, ordered by realised rate. *The coupling node HC12 fires at 0.98, confirming that schedule-driven cost is near-certain to be incurred once the chain is active. Source systems fire at their base rates.*

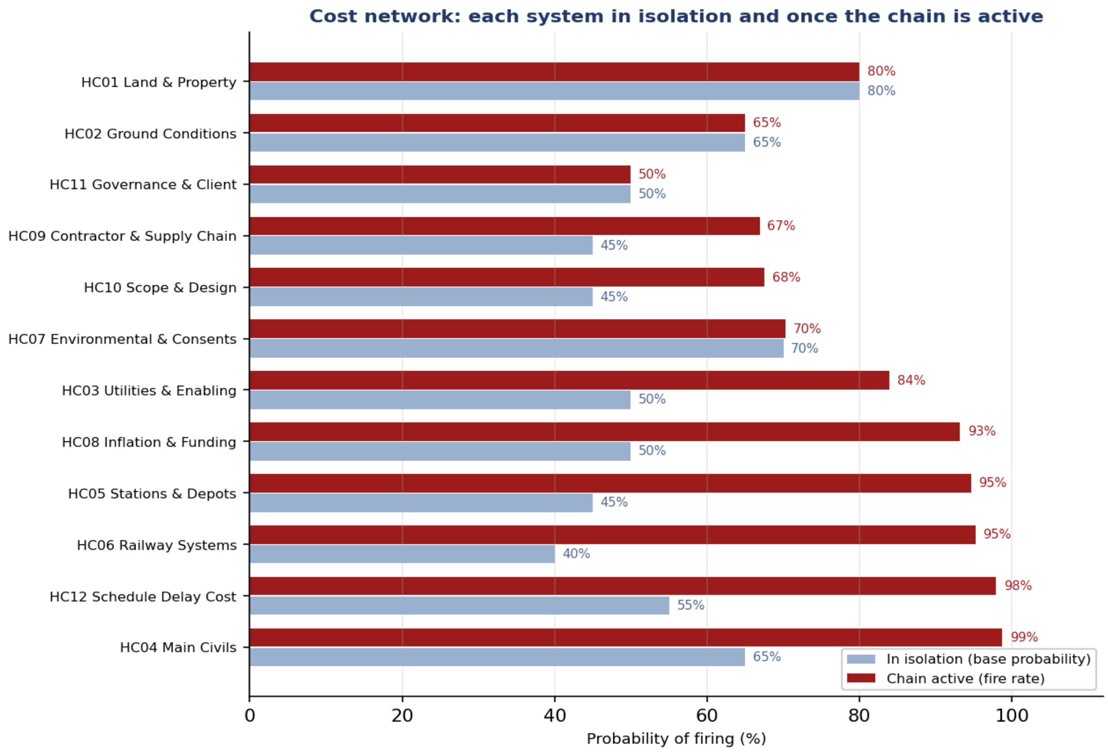
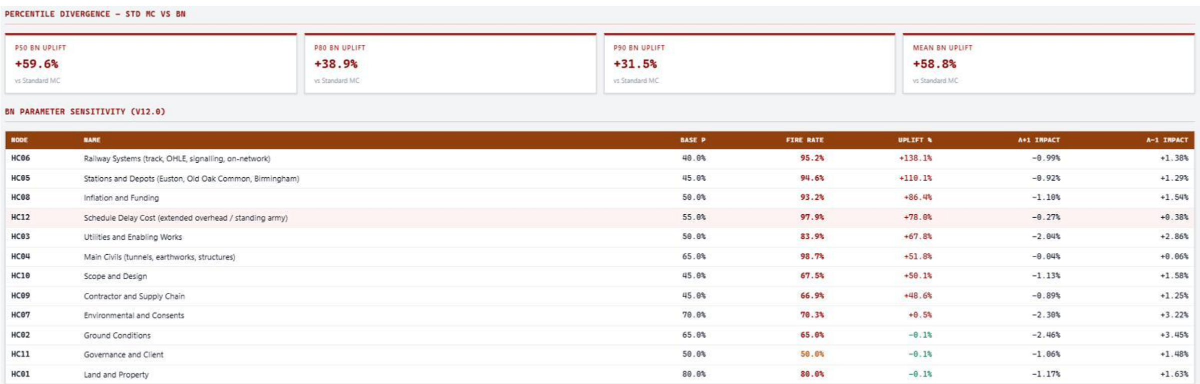


Figure 6a. The cost mechanism drawn: standalone probability against realised firing rate for the twelve cost systems, from the engine's node diagnostics. *The four source systems are unchanged; main civils rises from 0.65 to 0.987 and the coupling node HC12 from 0.55 to 0.979.*



PERCENTILE DIVERGENCE – STD MC VS BN

P50 BN UPLIFT	P80 BN UPLIFT	P90 BN UPLIFT	MEAN BN UPLIFT
+59.6%	+38.9%	+31.5%	+58.8%
vs Standard MC	vs Standard MC	vs Standard MC	vs Standard MC

BN PARAMETER SENSITIVITY (V12.0)

NODE	NAME	BASE P	FIRE RATE	UPLIFT %	A+1 IMPACT	A-1 IMPACT
HC06	Railway Systems (track, OHLE, signalling, on-network)	40.0%	95.2%	+138.1%	-0.99%	+1.38%
HC05	Stations and Depots (Euston, Old Oak Common, Birmingham)	45.0%	94.6%	+110.1%	-0.92%	+1.29%
HC08	Inflation and Funding	50.0%	93.2%	+86.4%	-1.10%	+1.54%
HC12	Schedule Delay Cost (extended overhead / standing army)	55.0%	97.9%	+78.0%	-0.27%	+0.38%
HC03	Utilities and Enabling Works	50.0%	83.9%	+67.8%	-2.04%	+2.86%
HC04	Main Civils (tunnels, earthworks, structures)	65.0%	98.7%	+51.8%	-0.04%	+0.06%
HC10	Scope and Design	45.0%	67.5%	+50.1%	-1.13%	+1.58%
HC09	Contractor and Supply Chain	45.0%	66.9%	+48.6%	-0.89%	+1.25%
HC07	Environmental and Consents	70.0%	70.3%	+0.5%	-2.30%	+3.22%
HC02	Ground Conditions	65.0%	65.0%	-0.1%	-2.46%	+3.45%
HC11	Governance and Client	50.0%	50.0%	-0.1%	-1.06%	+1.48%
HC01	Land and Property	80.0%	80.0%	-0.1%	-1.17%	+1.63%

Figure 6b. Percentile divergence on the cost network, from the engine's comparison dashboard*. The P50 rises by 59.6 per cent, the P90 by 31.5 per cent and the mean by 58.8 per cent. As on the schedule, the body of the distribution moves further than the tail: the chain relocates the centre rather than merely thickening the extreme.*

The redistribution of variance is more pronounced on cost than on schedule, and it is instructive. Table 9 ranks the variance contributors before and after the dependencies are switched on. Under independence, main civils dominates, contributing more than a third of the variance; the schedule-delay cost is second and inflation third. Under the chain, the order is overturned: inflation rises to first and main civils falls to second, its share roughly halving, from about 35 to 17 per cent. The reason is the same counter-intuitive mechanism seen on the schedule side, in stronger form: main civils now fires almost every time (0.99), so it has almost no variability left to contribute, and it recedes into the baseline. Meanwhile scope and contractor, which sit upstream and remain genuinely uncertain, more than double their shares. A dependency-aware model does not simply raise the total; it relocates attention from the system that is certain to be hit to the systems whose behaviour still swings the outcome, which is precisely where management effort has purchase.

System	Independent (var %)	Chain (var %)	Movement
Inflation and funding	14.6	18.2	rises to 1st
Main civils	34.7	17.0	falls to 2nd (near-certain)
Schedule delay cost (HC12)	21.5	16.9	now 3rd
Scope and design	6.8	14.8	more than doubles
Contractor and supply chain	6.5	13.7	more than doubles
Governance and client	3.1	6.3	doubles
Stations and depots	7.8	6.2	falls slightly
All other systems (combined)	5.0	6.9	residual

Table 9. Variance contribution of each cost system, independent versus dependency-aware. *The headline reversal (inflation overtaking main civils, and main civils halving) is the signature of a convergence node becoming near-certain. Contributions are the engine's regression-based variance shares (normalised squared standardised betas) and sum to 100% within rounding.*

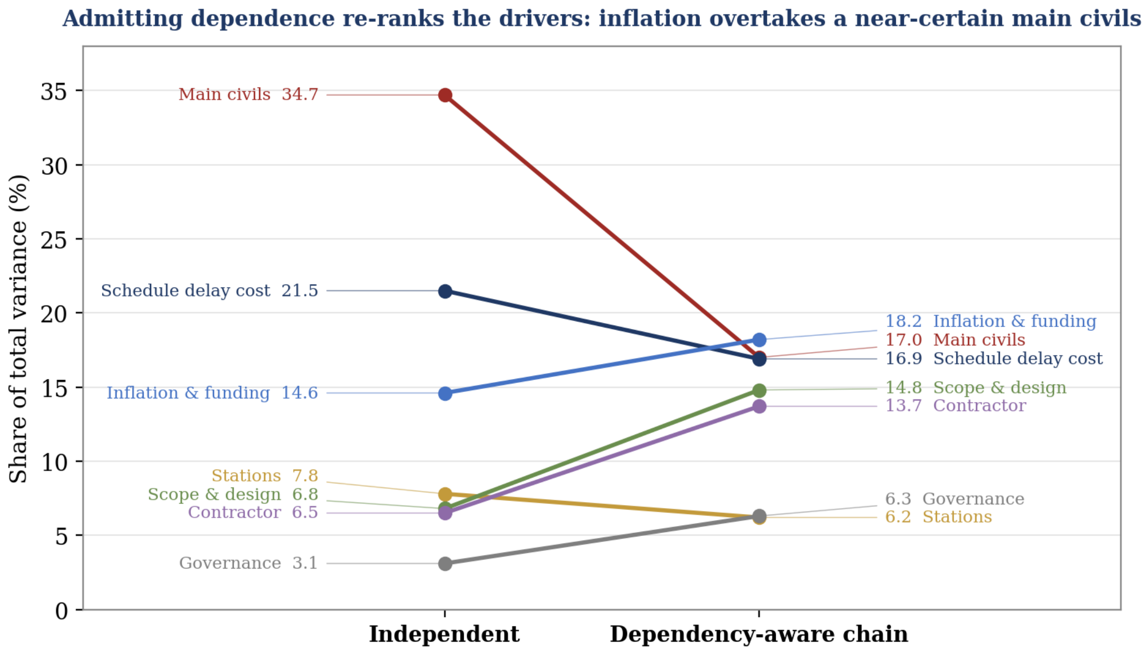


Figure 7. The variance decomposition of Table 9, drawn as a re-ranking. *Under independence main civils dominates at 34.7 per cent; under the chain it falls to 17.0 per cent (firing almost every run, it has little variability left to contribute) while inflation rises to first and scope and contractor more than double their shares.*

5.4 Lenses five and six: the cost updated at the 2026 endpoint

Lenses five and six take the dependency-aware chain of lens four and let it learn. The evidence is the public record itself, organised into the two observation registers of Section 4.4 and released with this paper: the *official* register, which records only what the record states, and the *modelled* register, the sponsor-grade magnitudes each anchored to a dated document; among them, the disclosed de-specification is entered as an explicit saving on HC10 (ITV News, 2026; New Civil Engineer, 2026). The updating is performed under the rolling-origin discipline of Section 5.7, each document admitted only at the cut-off on or after its publication date, and the lens values reported here are the 2026 endpoints of those two released runs. On cost, the official lens reaches a P90 of £91.8bn and the modelled lens £105.4bn; on schedule, 4,671 and 6,187 days. The official cost lens stands inside the reset band; the modelled lens carries about £2.7bn past its ceiling; both schedule lenses stand inside the ten-to-seventeen-year band.

The direction of travel is the finding. The chain reaches the floor of the band from the coupled register alone; the record then carries the figure upward as its magnitudes are disclosed, by about three billion on the official track and about seventeen on the modelled. The distance between the two updated lenses is the disclosure gap, measured directly: the record names the programme's risks but quantifies few of them, so a baseline fed only the record lands well short of one fed the magnitudes the sponsor already held. What disclosure adds is not a new level but the remainder of a level the structure had already set.

Cost lens (£bn)	P50	P80	P90	vs reset £87.7–102.7bn
Independent (2013-style)	46.4	60.2	67.4	below the band
Dependency-aware (chain)	73.7	83.7	88.6	P90 at band floor
Updated 2026; official (forward log)	77.5	87.1	91.8	inside the band
Updated 2026; modelled (forward log)	**88.9**	**100.0**	**105.4**	**above ceiling by £2.7bn**

Table 10. The four cost lenses against the 2026 reset range. *The updated rows are the 2026 endpoints of the rolling-origin runs of Section 5.7, transcribed from the released Intelligence Logs, which publish P50, P80 and P90; the five-seed tail averages for these lenses are given in Section 6.10. The chain reaches the floor of the band from the coupled register alone; the official record carries the P90 to £91.8bn, and the sponsor-grade magnitudes to £105.4bn. The robust, magnitude-independent quantity remains the £21.3bn independence gap of Section 5.3.*

A caveat of interpretation applies to the chain specifically, and is stated here rather than buried. Unlike a frozen-in-2013 forecast, the register magnitudes in Table 3 are anchored to a mixture of the programme's 2012–2013 base documents and to later disclosures; the November 2023 main-civils disaggregation, and, through HC12, the 2026 reset slip. The chain's reaching the reset floor is therefore not a claim of pure ex-ante foresight; the magnitudes incorporate information that emerged after 2013. What is genuinely structure-only, and independent of those later anchors, is the £21.3bn independence gap of Section 5.3, because that figure is the difference between two runs on identical magnitudes. The distinction between the structural gap (magnitude-independent) and the absolute level (magnitude-dependent) is important enough that Section 7.4 records it as a correction to the framing used in earlier drafts of this work.

5.5 Where a mitigation pound earns most: a what-if on the arrows

A dependency-aware model carries information an independent model cannot: because the systems are connected, weakening one connection propagates. This permits a what-if analysis: asking which single dependency, if it could be broken, would most reduce the P90. The results below are presented strictly as model what-ifs. The direction of each finding is robust, because it follows from the graph structure; the precise pound or day value attached to it is model-dependent, because it follows from the magnitudes. They identify where to look, not how much is guaranteed.

On the schedule network, the single most effective link to weaken is the one from funding instability to contractor performance (HS10→HS07). Cutting it in the what-if reduces the P90 slip from 4,535 days to about 3,718 days; a saving of roughly 817 days, about 2.2 years. The result is intuitive once the structure is seen: contractor performance is the convergence node, and funding instability is one of its strongest and most upstream drivers, so stabilising the funding profile relieves the system on which the most arrows converge. Figure 8 ranks the candidate links.

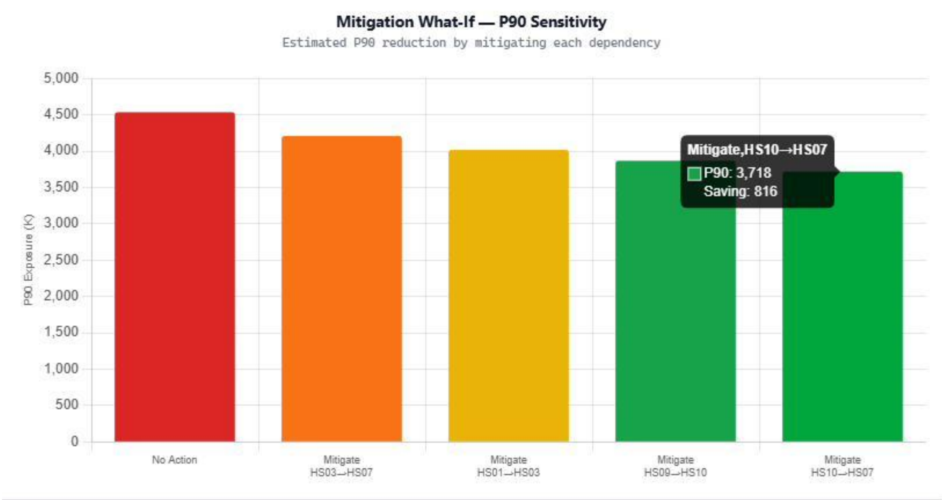


Figure 8. Schedule mitigation what-if: P90 programme slip under no action versus weakening each of the strongest single dependencies. *Breaking the funding-to-contractor link (HS10→HS07) yields the largest single reduction, about 816 days. Values are model what-ifs; the ranking reflects graph structure.*

On the cost network, the strongest single links cluster, as the structure predicts, around the main-civils convergence zone and its feeders. The most effective single link in the what-if is the one from scope and design to stations (HC10→HC05), whose weakening reduces the P90 from £88.6bn to about £72.7bn, a reduction of roughly £16bn. As with the schedule, the lesson is not the exact figure but the location: the purchase sits on the arrows feeding the most-connected nodes, not necessarily on the largest cost system in isolation. An independent model, lacking arrows, cannot produce this ranking at all; it can only rank systems by their standalone size, which, as Section 5.3 showed, is precisely the ranking that the dependence structure overturns.

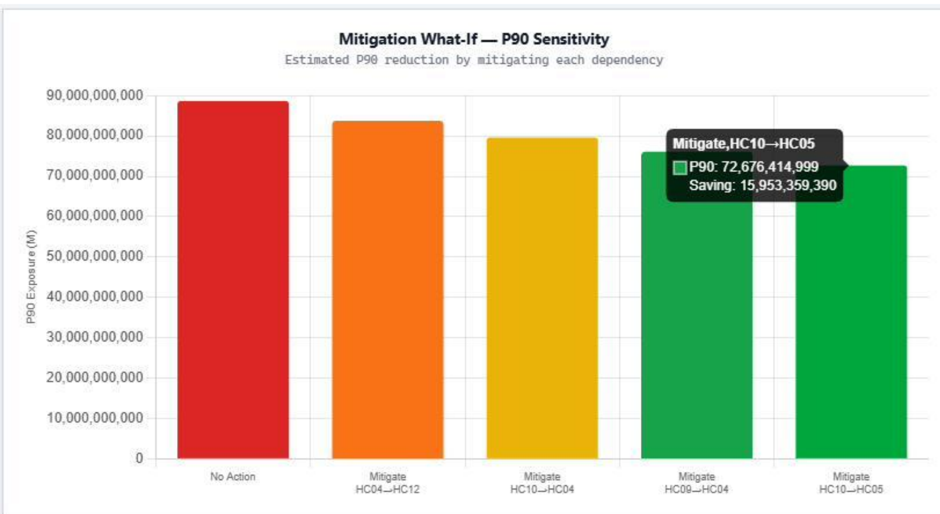


Figure 9. Cost mitigation what-if: P90 total cost under no action versus weakening each of the strongest single dependencies. *Breaking the scope-to-stations link (HC10→HC05)*

yields the largest single reduction, about £16bn. Values are model what-ifs; the ranking reflects graph structure.

5.6 The six lenses together

Table 11 draws the six lenses into a single view. Read down the schedule pair and then the cost quartet, two facts stand out. First, in each domain the move from an independent to a dependency-aware model, with nothing else changed, is the largest single step in the table: 864 days on schedule and £21.3bn on cost, the independence error itself. Second, the evidence then moves the figure in one direction only, upward, and by an amount that depends on what is disclosed: the public record adds about three billion to the cost P90, and the sponsor-grade magnitudes about seventeen, carrying the modelled lens about £2.7bn past the band ceiling. The structure sets the level; disclosure completes it.

#	Domain	Lens (configuration)	P90	Position
1	Schedule	Independent (2013-style)	3,671 d (10.1 yr)	baseline
2	Schedule	Dependency-aware (chain)	4,535 d (12.4 yr)	+864 d gap; in reset band
3	Cost	Independent with correlation Iman Conover (2013-style)	£67.4bn	baseline
4	Cost	Dependency-aware (chain)	£88.6bn	+£21.3bn gap; at reset floor
5	Cost	Updated 2026; official (forward log)	£91.8bn	inside the band
6	**Cost**	**Updated 2026; modelled (forward log)**	**£105.4bn**	**above ceiling by £2.7bn**

***Table 11. The six lenses summarised**. The independence error (the gap between the independent and dependency-aware lens in each domain) is 864 days on schedule and £21.3bn on cost. On schedule the three dependency-aware lenses span the ten-to-seventeen-year band; on cost the chain stands at the floor, the official lens inside the band, and the modelled lens about £2.7bn above its ceiling.*

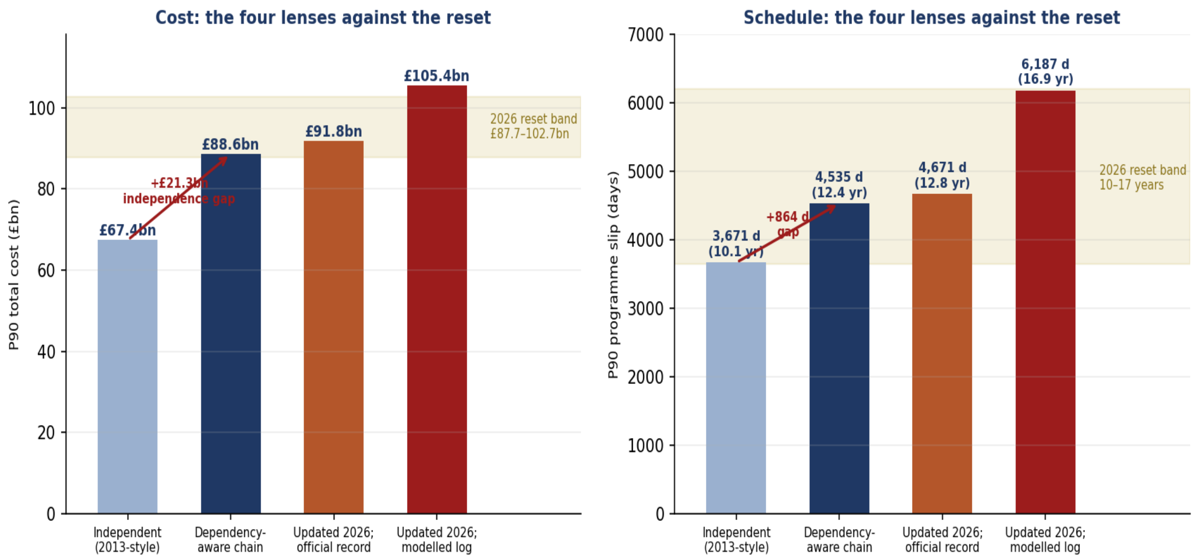


Figure 10. The four lenses in each domain against the programme's 2026 reset. *Left: cost P90 by lens against the reset band of £87.7–102.7bn; the chain reaches the floor from the coupled register alone, the official record carries the figure to £91.8bn and the sponsor-grade magnitudes to £105.4bn. Right: the schedule counterparts against the ten-to-seventeen-year band. The red arrow in each panel is the independence gap of Section 5.2.*

5.7 The rolling-origin forward test

Lenses five and six report a destination; this section reports the route, because the endpoint is produced by the walk itself. The baseline is carried forward one cut-off at a time, with each figure admitted only once the document that carries it is dated. That is the rolling-origin test, and it has now been run: for both domains, and on two separated evidence tracks. The *official* track admits only what the public record quantifies; the *modelled* track admits the disclosed magnitudes a sponsor holds, each on its publication date. Neither track sees a figure before its source exists. The schedule of what enters when is given in the companion provenance dossier.

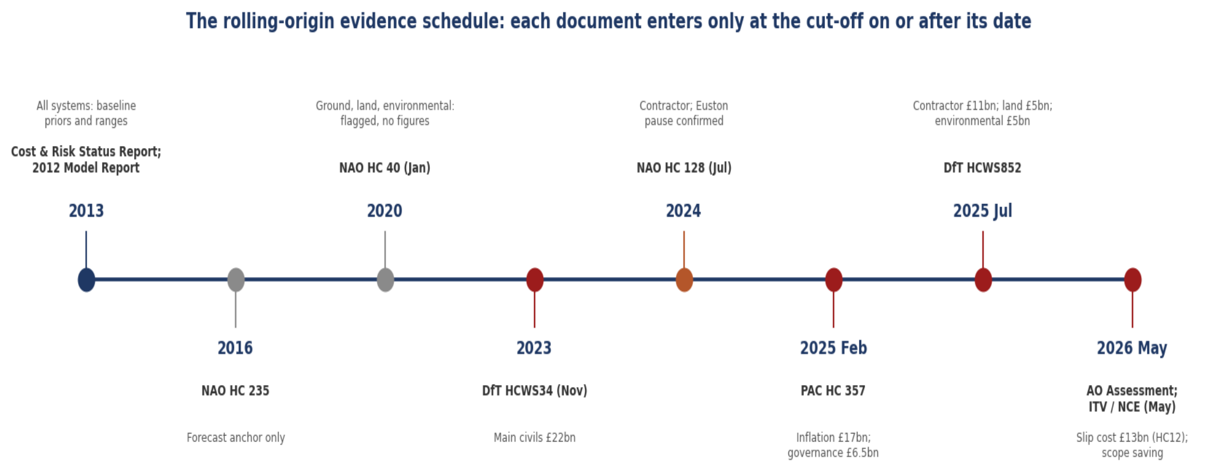


***Figure 10a. The rolling-origin evidence schedule**. Each public document enters the log only at the cut-off on or after its publication date; the largest sourced magnitude, the £22bn main-civils figure of November 2023, enters two and a half years before the official reset absorbed it.*

The result is the same in both domains, and it shows in motion what the endpoint lenses of Section 5.4 report at rest. Each track begins, at 2013, exactly on the structural chain (£88.6bn on cost, 4,535 days on schedule) confirming that the forward machinery reproduces the chain when given nothing but the 2013 register. Each then *falls* at the 2020 cut-off, cost to about £86.6bn and schedule to about 4,340 days, before rising as later magnitudes are disclosed. That fall is not incidental; it is the proof of independence. A model that accumulated state across cut-offs could only ever climb, because evidence would pile on evidence; a model that resets to a fixed prior before each cut-off can fall when the evidence of a given year is, on net, mild. The dip appears in all four tracks, which is four independent confirmations that each cut-off runs from the same 2013 baseline rather than from the residue of the one before.

Cut-off	Cost (official	Cost) modelled	Schedule (official	Schedule) modelled
2013	£88.6bn	£88.6bn	4,535 d	4,535 d
2020	£86.6bn	£86.5bn	4,336 d	4,360 d
2023	£86.6bn	£90.1bn	4,599 d	5,178 d
2024	£89.4bn	£98.6bn	4,672 d	6,157 d
2025 Jul	£91.3bn	£103.2bn	4,685 d	6,184 d
2026	**£91.8bn**	**£105.4bn**	**4,671 d**	**6,187 d**

***Table 12. The rolling-origin forward test:** P90 at each cut-off, by domain and evidence track, from the released Intelligence Logs. All four tracks reproduce the 2013 chain baseline, fall at the 2020 cut-off (the independence signature), and rise as magnitudes are disclosed. The endpoint is reached only when the last large magnitude (the schedule slip, in May 2026) is admitted.*

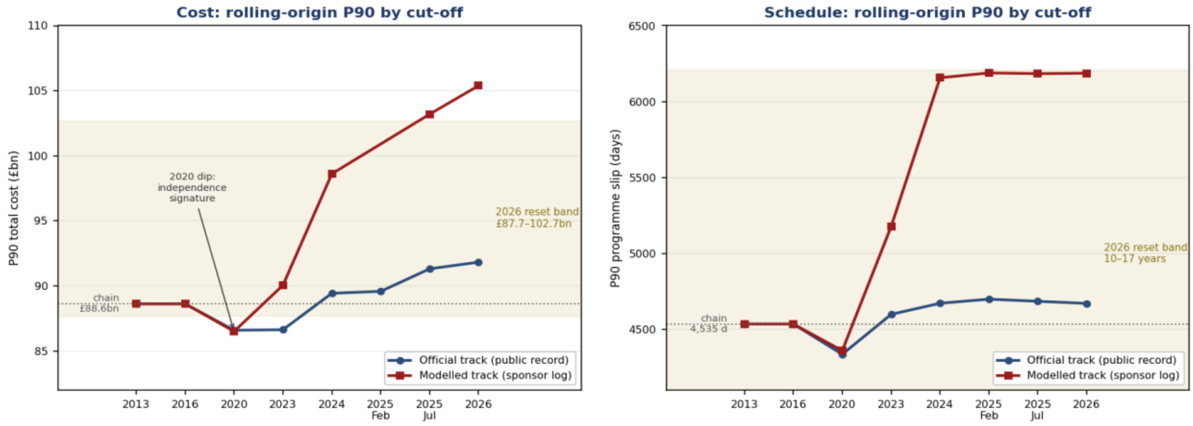


Figure 10b. The four rolling-origin tracks against the 2026 reset. *In each domain the modelled track (which admits disclosed magnitudes on their publication dates) climbs into the reset band and past its ceiling on cost, while the official track (which admits only what the record quantifies) lands well short. The gap between the two tracks is the disclosure gap, measured directly.*

The divergence between the tracks is the finding. The modelled track reaches £105.4bn on cost and 6,187 days on schedule; inside the reset band on schedule, and about £2.7bn above its ceiling on cost; the official track reaches only £91.8bn and 4,671 days. The distance between them is the disclosure gap, now measured rather than asserted: the public record named the programme's risks but quantified few of them, so a baseline fed only that record lands far short of one fed the magnitudes the sponsor already held. The endpoint values of Section 5.4 are the last row of Table 12; what the trajectory adds is the timing. The record carried the probabilities early and the magnitudes late, and it is the magnitudes, disclosed piecemeal across a decade, that carry the P90 into the reset band. The answer was assemblable from public documents the whole time. What was missing was a structure to sum them and a baseline to remember them as they arrived.

Two cautions bound the claim, and neither is small. The graph was drawn in 2026, with the outcome known, so a frozen set of documents is not the same as a mind frozen in 2013; the forward test disciplines the *information* available at each step, not the analyst's knowledge of where the account ends. And the modelled magnitudes, though each is anchored to a dated document, were selected knowing which risks proved material. The forward test therefore demonstrates that a baseline with memory, walked forward under strict information cut-offs, would have tracked the disclosures as they were published; the £22bn main-civils figure entering in November 2023, two and a half years before the reset absorbed it. It does not, and cannot on a completed project, demonstrate prospective accuracy on an outcome that is not yet known. That step is reserved for a programme still in flight, and is the subject of Section 7.4.

5.8 Two mitigation lists: the tornado and the chain

Every risk process ends in the same meeting: a finite mitigation budget and the question of where it goes. A standard Monte Carlo model answers with a tornado, the systems ranked by their standalone contribution to output variance, and on this register that ranking is unambiguous. The independent model's top five are main civils at 34.7 per cent, the schedule-delay cost at 21.5, inflation and funding at 14.6, stations at 7.8 and scope and design at 6.8 (Table 9, and the left panel of Figure 10c). It is a defensible list, and it is the list a 2013-style model hands the board.

The same register with its arrows switched on returns a different answer twice over. The first correction is the re-ranking already reported in Section 5.3: inflation rises to first, scope and design and the contractor more than double their shares, and main civils halves (Figure 7, and the right panel of Figure 10c). The second correction is the deeper one. Four of the tornado's five are, under the chain, near-certainties: main civils fires at 0.987, the schedule-delay cost at 0.979, stations at 0.946 and inflation at 0.932 (Table 8). Their tornado bars are not intrinsic properties of those systems; they are inherited through the arrows that feed them. A tornado drawn under independence therefore ranks pass-throughs. It measures where the loss will be booked, not where it is generated.

The arrows of Section 4.2 say where it is generated, because they form one legible spine. Governance instability raises the probability of scope and design change (HC11 to HC10); scope change loads main civils and the stations in parallel (HC10 to HC04 and HC10 to HC05); civils drive the schedule-delay cost (HC04 to HC12); delay drives the inflation and funding exposure (HC12 to HC08); and stations and inflation together load railway systems (HC05 and HC08 to HC06). Read against the tornado, the comparison is stark: four of the tornado's top five (civils, delay cost, inflation, stations) are links of this one chain, while the chain's driver, scope and design, sits fifth at 6.8 per cent and its origin, governance, sits near the bottom at 3.1. The what-if of Section 5.5 prices the difference. Weakening the single mid-chain arrow from scope to stations (HC10 to HC05) returns the P90 from £88.6bn to about £72.7bn: roughly £16bn, about three-quarters of the entire £21.3bn independence gap, from one connection the tornado cannot see because an independent model has no arrows to rank. The schedule network tells the same story in days: its spine runs from governance through funding instability into the contractor convergence (HS09 to HS10 to HS07), and cutting the single funding-to-contractor arrow returns 817 of the 864-day gap, about 95 per cent.

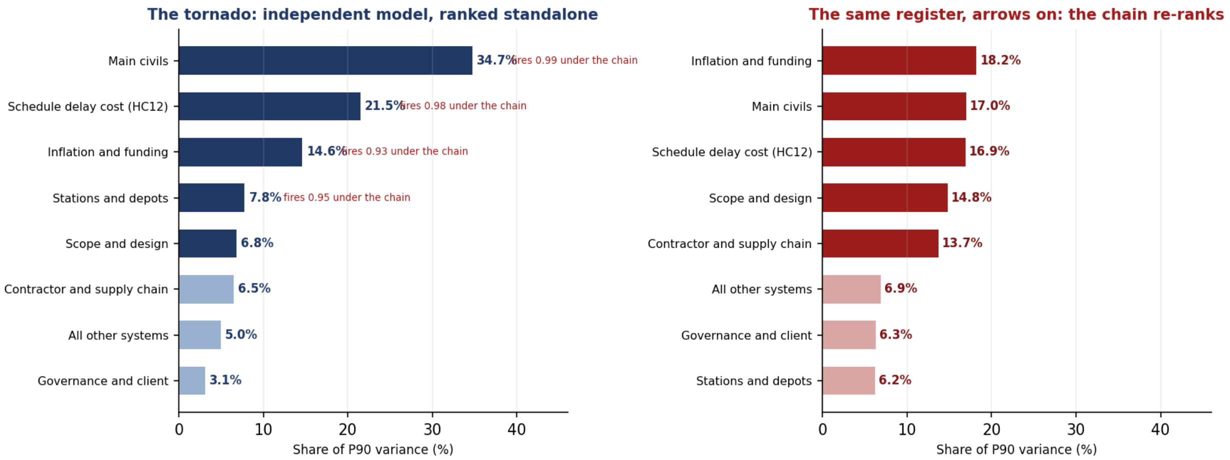


Figure 10c. The two mitigation lists drawn from one register*. Left: the standard tornado under independence, systems ranked by standalone variance share (top five in dark), with each top system's realised firing rate once the chain is active annotated in red. Right: the same shares with the arrows on. Shares are the regression-based contributions of Table 9; firing rates are the node diagnostics of Table 8; the priced interventions are the what-ifs of Section 5.5.*

The management reading follows directly, and it is the practical case for the coupled model. A pound aimed at the tornado's top bar buys magnitude relief on a system that the chain will set alight anyway; a pound aimed at the conditional, whatever instrument weakens the probability that a fired parent ignites its child, in contracting, in interface management, in sequencing, acts on the mechanism that gives the top bars their height. The two lists agree on almost nothing above the fold, and only one of them can be tested by the register itself: cut the arrow in the model, re-run, and read the price. These what-ifs remain model statements, as Section 5.5 states, and their magnitudes are indicative rather than engineering estimates; but their direction follows from the structure, and the structure is documented arrow by arrow. An independent model cannot produce this list at all. That is not a nuance of presentation. It is a different answer to where the next pound goes.

6. Validation and internal audit

A model that produces a striking number is only as good as the checks that surround it. This section reports, first, five internal checks (that each simulation has converged; that it is exactly reproducible; that its headline mean can be reconstructed from first principles, independently of the engine that produced it; that the correlation stage contributes essentially nothing; and that the independence gap is stable under resampling) and then three robustness tests that attack the gap directly: a structural-robustness sweep (the topology test), a distributional-shape test, and a per-node component-breach test. Each check was performed on the raw output by a separate routine, not read back from the engine's own summary, and every reconstruction passes.

6.1 Convergence

Every run in the study was executed at 100,000 iterations, and the percentiles are converged. On the cost network the chain P90 reads £88.64bn at 35,000 iterations, £88.67bn at 50,000 and £88.65bn at 100,000; a drift under five-hundredths of one per cent beyond 35,000. On the schedule network the P90 reads 4,535 days at 35,000, 4,536 at 50,000 and 4,532 at 100,000, the same picture. The median and P80 are equally flat in both domains. The headline figures are therefore properties of the models, not of the sample size, and a reader re-running at, say, 50,000 iterations would obtain the same answer to the reported precision.

6.2 Reproducibility and the engine record

Each configuration carries an input digest and a configuration hash, so that a run can be reproduced and verified.

6.3 Reconstruction of the mean from first principles

The most informative check recomputes the expected total by hand and compares it with the engine. For each system, the expected contribution is its realised firing rate multiplied by the mean of its PERT magnitude, where the PERT mean is (min + 4×most-likely + max) / 6 at the shape used. Summing these across the twelve cost systems gives £73.747bn, against the engine's reported mean of £73.745bn; a difference of £0.002bn, under one part in thirty thousand. The same reconstruction on the schedule network gives 3,510.8 days against the engine's 3,512.1, a difference of 1.3 days. The engine is doing exactly the arithmetic the model specifies, and the headline magnitudes are not an artefact of any hidden step. More broadly, recomputing the P90 and the mean directly from each run's raw output vector reproduces every reported figure in this paper to the penny, across all twenty runs.

6.4 The correlation stage contributes essentially nothing

Section 3.3 distinguished two channels of dependence (the structural cascade and the engine's residual rank-correlation stage) and claimed the result rests on the former. This is confirmed by switching the correlation stage off and on. On the independent cost network the P90 with the correlation stage off is £67.373bn; with it on, £67.376bn (a difference of about £0.003bn. On the schedule network the same switch moves the independent P90 from 3,665 to 3,671 days, a difference of about six days. In each domain the correlation channel accounts for well under one per cent of the independence gap (£21.3bn on cost, 864 days on schedule). The dependence effect of Section 5 therefore cannot be attributed to a tuned correlation coefficient; it arises essentially wholly from the documented causal arrows. A complete separation) the full chain run with the correlation stage removed; is listed in Section 7.4, but the channel it would isolate is, on this evidence, too small to alter any conclusion.

6.5 The independence gap is sampling-stable

Finally, the £21.3bn cost gap is not a fragile difference between two noisy tails. Resampling the raw output of both the independent and the dependency-aware runs and recomputing the P90 difference many times places the bootstrap standard error at about £0.10bn, around a mean of £21.26bn; the same procedure on the schedule gap gives 864.5 days with a standard error of about 7 days. Each gap is large relative to its sampling error by roughly two orders of magnitude. The headline independence errors are structural properties of the models, reproducible and stable, not sampling accidents.

Check	What is compared	Result	Status
Raw-output reconstruction	reported P90 & mean vs raw vector, all 20 runs	reproduced exactly (Δ = 0)	pass
Mean reconstruction (cost)	hand-computed vs engine mean	£73.747bn vs £73.745bn (Δ £0.002bn)	pass
Mean reconstruction (schedule)	hand-computed vs engine mean	3,510.8 vs 3,512.1 days (Δ 1.a3 d)	pass
Correlation stage (cost P90)	stage off vs on (independent)	Δ £0.003bn	pass (negligible)
Convergence (cost P90)	P90 at 35k vs 100k iterations	£88.64bn vs £88.65bn	pass (< 0.05% drift)
Independence-gap stability	bootstrap SE on the cost P90 gap	£21.26bn ± £0.10bn	pass

***Table 13. Internal audit of the headline figures**. Each check was computed independently of the engine's own summary, on the raw output. The structural-robustness sweep, the distributional test and the component-breach test follow in Sections 6.6–6.8; the decomposition of that effect into its conditional-structure and noisy-OR components is reported in Section 6.5.*

6.6 Decomposing the dependence effect: structure and combination rule

Switching on the dependence does two things at once, and a rigorous account separates them. The conditional structure (the arrows) changes how often each system fires, because a system that fires raises its children's firing probability within the same run. The noisy-OR rule then governs how several fired parents combine at a shared child: each fired parent takes its own independent opportunity to set the child off, so a convergence node with several active parents fires more often than any single parent implies. To separate the two, the chain was re-run with the noisy-OR rule disabled; the arrows still active, but each child responding to its parents under a simple combination rather than the multiplicative rule. The difference between the independent model and this structure-only run is the work done by the conditional structure; the difference between the structure-only run and the full chain is the work done by the noisy-OR rule. Table 14 reports both, for both networks, at 100,000 iterations.

Configuration	Cost P90 (£bn)	Cost change	Schedule P90 (days)	Schedule change
Independent (no arrows)	67.4	baseline	3,671	baseline
+ conditional structure (simple combination)	83.0	+15.7 (74%)	4,180	+509 (59%)
+ noisy-OR rule (full chain)	**88.6**	**+5.6 (26%)**	**4,535**	**+355 (41%)**
Total independence gap		+21.3		+864

Table 14. The independence gap decomposed into the conditional structure and the noisy-OR combination rule, for both networks. *The conditional structure carries the majority of each gap on its own; the noisy-OR rule supplies the remainder.*

The result is the same in both domains, and it settles the most substantial objection to the method. The conditional structure carries the majority of the gap on its own, 74 per cent of the cost gap (£15.7bn of £21.3bn) and 59 per cent of the schedule gap (509 of 864 days), under a combination rule that does nothing more than let the arrows carry. The noisy-OR rule adds the remainder: 26 per cent on cost (£5.6bn) and 41 per cent on schedule (355 days). The dependence effect is therefore not an artefact of the particular combination rule chosen: most of it is the bare fact that the risks are connected, and would appear under any rule that propagated the connection at all. The noisy-OR contribution is real, and proportionally larger on the schedule side, where the convergence nodes carry more active parents; but it is the second term, not the first. Figure 11 shows the cost decomposition as a waterfall.

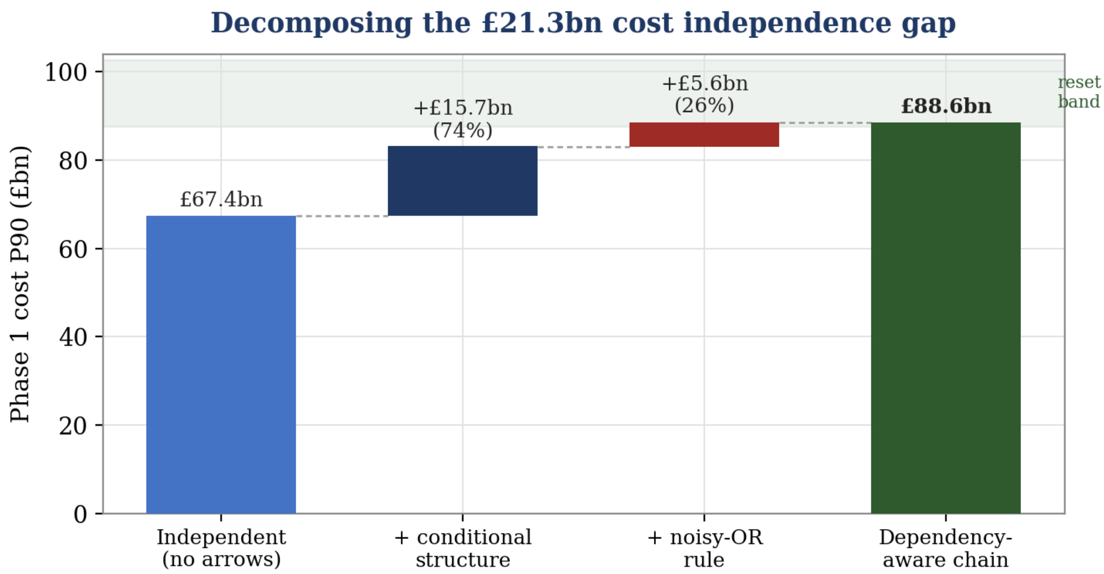


Figure 11. The £21.3bn cost independence gap decomposed. *From the independent P90 of £67.4bn, the conditional structure alone carries the figure to £83.0bn (+£15.7bn, 74 per cent of the gap); the noisy-OR combination rule then carries it to the full-chain P90 of £88.6bn (+£5.6bn, 26 per cent). The shaded band is the official reset range.*

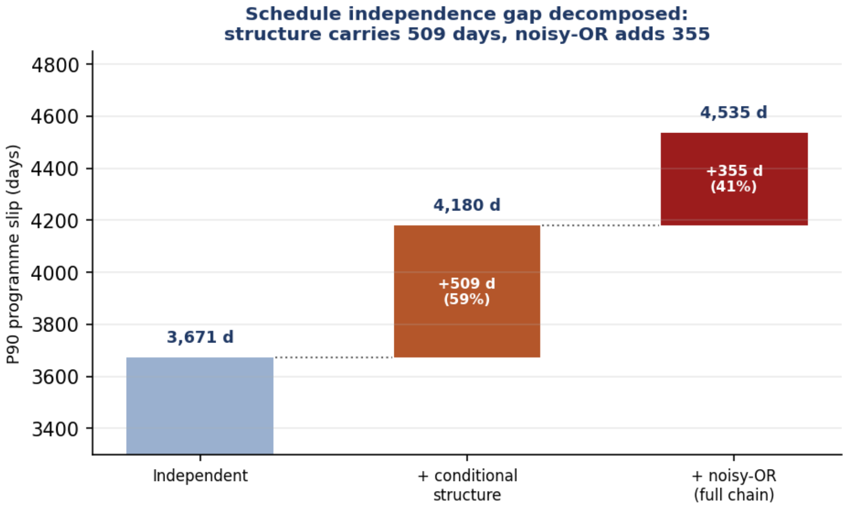


Figure 11a. The 864-day schedule independence gap decomposed in the same way: *from the independent 3,671 days, the conditional structure alone carries 509 days (59 per cent) and the noisy-OR rule the remaining 355 days (41 per cent), reaching the full-chain 4,535.*

6.7 Distributional robustness: the lognormal re-run

The dependency-aware distributions are close to symmetric, which invites the objection that the cascade has quietly suppressed a tail the magnitude shape would otherwise produce. To test this, each three-point PERT range was replaced by a lognormal magnitude matched to it in mean and variance, system by system, with the probabilities and the structure unchanged. The headline barely moves: the cost P90 goes from £88.63bn to £88.64bn, and the schedule P90 from 4,535 to 4,536 days, both within a day or a hundredth of a billion. What does change is the shape: the cost skewness turns from −0.05 to +0.13, and the schedule from +0.04 to +0.12, so a right tail reappears when the magnitudes are right-skewed; on the schedule side, where an independent lognormal run is also available, the gap holds under the new shape as well (an independent lognormal P90 of 3,667 days against the chain's 4,536, a gap of 869 days). Two conclusions follow, and both strengthen the argument: the independence gap is not an artefact of the PERT shape, since changing the shape leaves it intact; and the near-symmetry of the base runs is a property of the magnitude inputs, not of the cascade, since supplied with a right tail the chain carries one. The tail shape is an input choice, and it is labelled as such.

6.8 Structural robustness: the topology test

The most obvious objection to a dependency result is that the arrows were drawn to suit the conclusion. The answer is to redraw them. Four alternative graphs were constructed for each network, each a defensible reading of the same project: T2 reverses a salient edge (on cost, the schedule-cost-to-inflation link; on schedule, the governance-to-funding link); T3 adds a common-cause edge on top of T2; T4 is a deliberately sparse graph with two links removed; and T5 a deliberately dense graph with two further links added. All eight variants are acyclic, and each was run and compared with its independent baseline on identical magnitudes. Table 15 reports the independence gap under every variant.

Graph variant	Cost independence gap (P90, £bn)	Schedule independence gap (P90, days)
Base graph (T1)	**21.3**	**864**
T2; salient edge reversed	17.6	855
T3; common-cause edge added	17.7	895
T4; sparse (two edges removed)	19.3	833
T5; dense (two edges added)	21.7	891

Table 15. The independence gap under five graph shapes, each compared with its independent baseline on identical magnitudes. *The gap never closes: it sits between £17.6bn and £21.7bn on cost, and between 833 and 895 days on schedule. There is no reasonable structure in which independence is the cautious choice.*

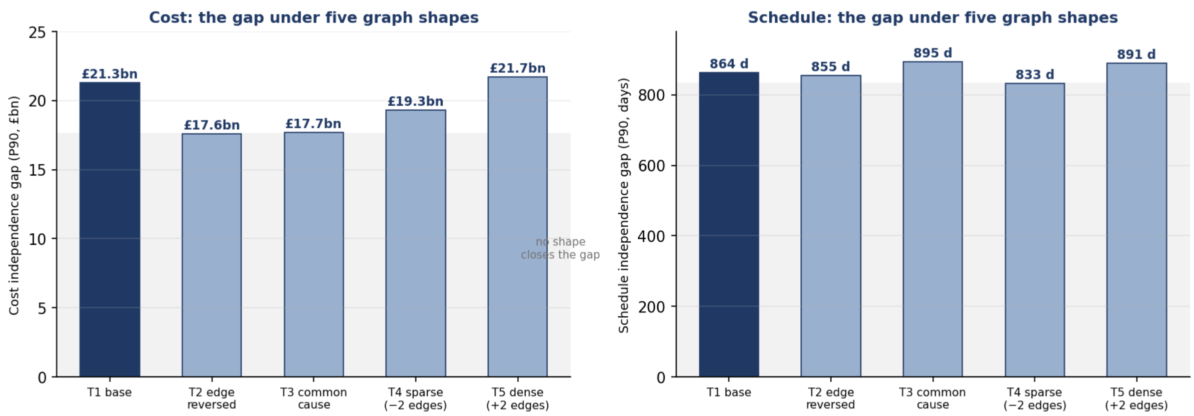


Figure 12. The independence gap under five graph shapes, cost left and schedule right (T1 base*; T2 salient edge reversed; T3 common-cause edge added; T4 sparse, two edges removed; T5 dense, two edges added). No defensible reading of the project closes the gap: the floor across shapes is £17.6bn on cost and 833 days on schedule, against base-graph values of £21.3bn and 864.*

The gap moves with the structure, as any honest model does (it is smallest under the reversed and common-cause cost variants and largest under the dense one) but it never closes, and across every variant the dependency-aware P90 stands well above the independent one. Which arrow to cut for mitigation depends on the structure; whether independence understates the total does not.

6.9 The component-breach test

The sharpest test is the one the programme itself ran, recorded in Table 2: every major Phase 1 component breached its own stated P90. The model can be asked how probable that pattern is, and the question is answered on the per-node output rather than argued. For each of the six driver systems, its own independent P90 ceiling is read from the independent-by-node run; the modelled-by-node run is then asked how often each system, and how many simultaneously, exceed those ceilings. Under independence each driver breaches its own P90 ten per cent of the time, by construction. Under the dependency-aware modelled run the breach rates rise to between 1.2 and 2.0 times that, and (the property that matters most) the number of simultaneous breaches gains both expectation and variance.

Cost driver	Indep. P90 ceiling (£bn)	Breach rate, modelled	Lift vs independent
Stations and depots	10.5	20.4%	2.0×
Inflation and funding	14.7	18.0%	1.8×
Schedule delay cost (HC12)	17.6	17.4%	1.7×
Main civils	22.9	14.9%	1.5×
Scope and design	10.1	13.1%	1.3×
Contractor and supply chain	9.8	12.0%	1.2×

Table 16. Per-driver breach of the independent P90 ceiling, cost network*. Ceilings are the independent-by-node P90; breach rates are counted in the modelled-by-node run. Every driver breaches its own independent ceiling more often once the dependence is active. Appendix D gives the full table for both networks.*

The joint behaviour is the point. Among the six cost drivers the expected number of simultaneous breaches rises from 0.600 under independence to 0.958 under the modelled chain, and its variance from 0.539 to 0.804. That variance exceeds the 0.54 that six independent one-in-ten events would produce: the breach count is over-dispersed, which is the statistical signature of co-movement (the systems breach together more than chance allows. The all-six joint breach, essentially impossible under independence (none in ten thousand runs, against the one-in-a-million an independent model implies), rises to about one in fifty thousand under the dependency-aware run; four-or-more rises from 0.13 to 0.77 per cent. Running the identical test against the per-node output of the pure dependency-aware chain) now available, rather than the evidence-revised modelled run; gives a larger lift still, the expected breach count reaching 1.03 and its variance 0.85, so the modelled figures are the conservative reading. The schedule network behaves the same way, more mildly: the expected breach count rises from 0.600 to 0.817 and its variance from 0.538 to 0.697, with every driver lifting except governance, whose prior the evidence revised downward (Section 5.4) so that it breaches slightly less.

Breach-count statistic (six drivers)	Cost: indep → modelled	Schedule: indep → modelled
Expected number of simultaneous breaches	0.600 → 0.958	0.600 → 0.817
Variance of that number	**0.539 → 0.804**	**0.538 → 0.697**
All six breach together	~0 → 1 in 50,000	~0 → ~0
Four or more breach	0.13% → 0.77%	0.14% → 0.39%

Table 17. The distribution of simultaneous breaches, independent versus modelled, for the six drivers in each domain. *The dependence lifts both the mean and the variance of the breach count; the variance rise is the over-dispersion an independence assumption cannot represent.*

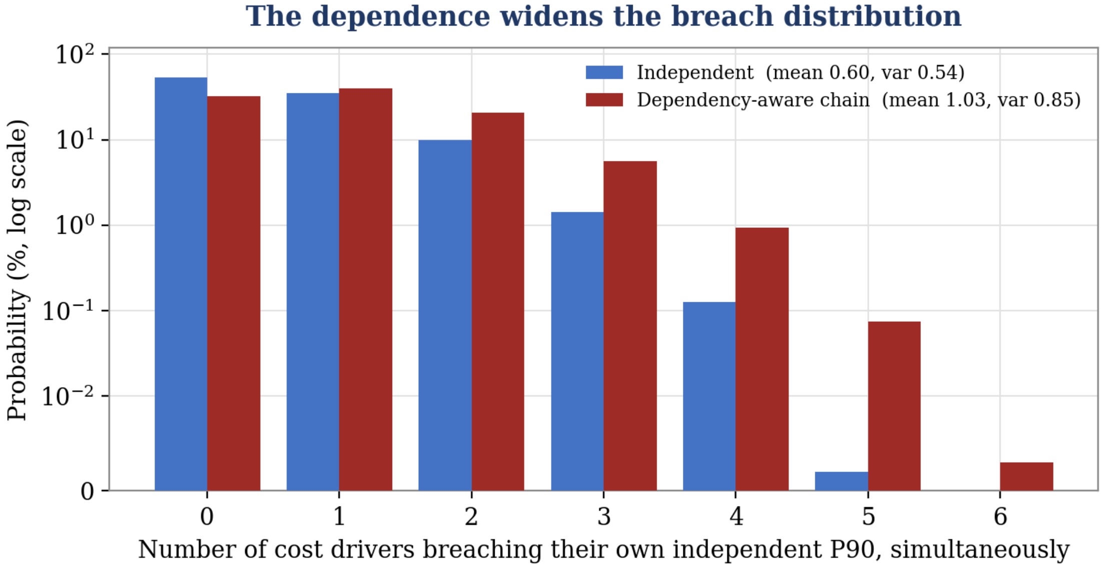


Figure 13. The distribution of simultaneous breaches among the six cost drivers, independent versus the dependency-aware chain (probability on a log scale). *The dependence shifts probability mass to the right and thickens the upper tail: the mean rises from 0.60 to 1.03 and the variance from 0.54 to 0.85; the over-dispersion an independence assumption cannot represent.*

HS2 delivered a breach on all six of its real components. Even the dependency-aware modelled run, carrying the downward belief revisions the public record forced, prices that joint outcome at roughly one in fifty thousand; so the realised event is more extreme still, and the independent model, which prices it at one in a million, was not being conservative; it was wrong in a way that resembled conservatism. What the dependence supplies is not certainty of disaster but the correct shape of the risk: a heavier, wider distribution of joint breaches than independence can express, lifting an outcome from effectively impossible to merely rare.

6.10 Multi-seed replication

The final robustness question is the seed. Every headline in this paper derives from runs at seed 4209; a sceptical reader is owed evidence that nothing turns on that choice. Five further independent seeds (1109, 2109, 3109, 4109 and 5109) were therefore run at 100,000 iterations each, across all four configurations and both networks: twenty runs, released as the multi-seed workbook alongside this paper. Table 18 reports the P90 by configuration.

Configuration	Schedule P90 (days), five seeds	Cost P90 (£bn), five seeds
Independent (2013-style)	3,671 ± 3 [3,668 – 3,677]	67.46 ± 0.06 [67.38 – 67.54]
Dependency-aware (chain)	4,537 ± 3 [4,534 – 4,541]	88.61 ± 0.02 [88.58 – 88.63]
Updated 2026; official	4,685 ± 9 [4,670 – 4,694]	92.02 ± 0.18 [91.81 – 92.30]
Updated 2026; modelled	6,221 ± 20 [6,206 – 6,261]	106.18 ± 0.09 [106.01 – 106.28]
Independence gap (chain − independent)	**865 ± 4 [860 – 872]**	**21.15 ± 0.06 [21.08 – 21.22]**

Table 18. The multi-seed study: *mean ± population standard deviation and range of the P90 across five independent seeds, by configuration and domain, from the released multi-seed workbook (100,000 iterations per run, PERT shape 4). The independence gap moves by less than half a per cent across seeds in either domain.*

Three readings follow. First, the gap is seed-stable to a degree that removes the question: a standard deviation of £0.06bn on a £21bn effect, and four days on an 864-day effect, is seed noise three orders of magnitude below the quantity being measured, and it agrees with the bootstrap standard error of Section 6.5. Second, the reporting seed is representative: its independent and chain values sit inside the five-seed range in both domains, at or adjacent to the range edges; its gap (£21.25bn) consequently sits three-hundredths of a billion above the five-seed span, the arithmetic consequence of those two components lying at opposite edges of their own ranges. Third, the two modelled endpoints of the reporting run sit slightly below the five-seed band, £105.4bn against 106.0 to 106.3, and 6,187 days against 6,206 to 6,261: a difference under one per cent, consistent with the session-level posterior draw of the MCMC stage, and conservative in direction, since the released endpoint understates rather than overstates the seed-study centre. The same five runs supply the tail measure the endpoint logs do not publish: across seeds, the CVaR at 90 per cent for the official configuration is about £97.1bn on cost and 5,043 days on schedule, and for the modelled configuration about £111.3bn and 6,591 days, so the updated lenses carry a tail average as well as a quantile. One transcription defect in an earlier issue of the workbook, in which the seed-5109 official-cost row duplicated the chain row because that run had not been executed, was caught by audit; the run was then executed and the corrected workbook is the one released.

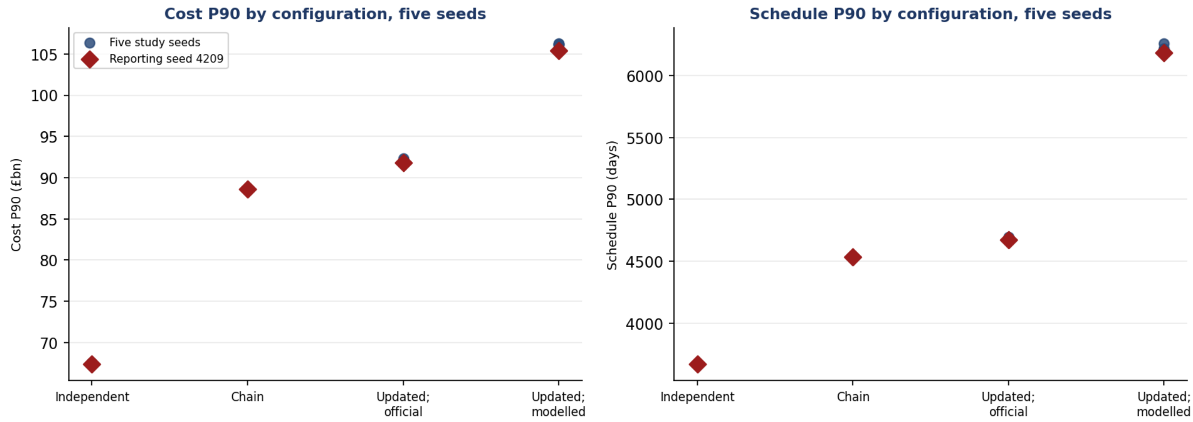


Figure 13a. The multi-seed study drawn: *P90 by configuration for the five study seeds (circles) with the reporting seed 4209 marked (diamonds), cost left and schedule right. The clusters are tight at the scale of the effects; the reporting seed sits within or marginally below each cluster.*

7. Discussion

7.1 Objections and responses

That independence is a conservative, safe assumption. It is not. Independence is a precise claim that the principal hazards cannot arrive together. On a construction programme that claim is false, and being false in the optimistic direction, it systematically understates risk. The £21.3bn and 864-day gaps measure exactly how much. A genuinely cautious assumption would err high; the independence assumption errs low, which is the opposite of caution.

That the result depends on hindsight. The two central quantities do not. The independence gap is the difference between two runs on identical magnitudes, distinguished only by whether the documented arrows are switched on. No outcome data enters that comparison; it is a property of the register's structure, computable in 2013. The reset band is used only as a target to check the updated lenses against, not as an input to the gap.

That the reference class already tells us the answer, so the model is redundant. The reference class identifies the likely destination; that a rail programme of this kind tends to overrun substantially (Flyvbjerg, Holm and Buhl, 2002). It does not identify the route: which systems drive the overrun, where the purchase sits, how schedule slip converts to cost. The network model supplies the route, and the route is what a sponsor can act on. The two are complementary: the reference class bounds the optimism, and the network explains the mechanism.

That the figures are simply invented. Every node magnitude is drawn from the programme's own registers (Tables 2–4), every dependency from documented engineering logic (Appendix C), and every headline figure is reconstructed from first principles in Section 6. The one frankly modelled quantity (the 13.6 per cent overhead share in the coupling) is labelled as such and bounded by a sourced floor. The structural gap, which is the paper's central claim, is independent of all magnitudes.

7.2 Transferability: the diagnostic any project can run

Nothing in the method is specific to HS2. Any project that maintains a quantitative risk register can compute its own independence error by running that register twice: once as it stands, and once with the documented dependencies between its risks switched on, holding magnitudes, seed and iteration count fixed. The difference between the two P90 values, in the project's own currency and its own units of time, is the cost its independence assumption is concealing. The marginal expense is one additional simulation, of the order of seconds of compute. A second, equally cheap practice follows: where a project couples its schedule and cost models through a single documented node, as here through HC12, it makes explicit (and therefore auditable) the channel by which programme slip becomes money, instead of leaving that channel to an assumed correlation or to none at all.

7.3 Recommendations for practice

Three recommendations follow directly from the results, and none requires new tooling or new data. First, run the register twice and report the independence error as a governance metric in its own right; a project that cannot state its independence error does not know how much of its risk its model is suppressing. Second, allow the baseline to learn: update node probabilities and magnitudes as evidence arrives, and report the updated distribution rather than defending the figure set at the moment of least information. Third, identify and watch the convergence nodes (the systems on which many arrows land, such as contractor performance on the schedule and main civils on the cost) because, as Section 5 showed, dependence relocates both the risk and the management purchase onto and around those nodes, away from the systems that look largest in isolation.

7.4 Limitations and remaining work

This paper establishes the method, the schedule-to-cost coupling and the priced independence gap, with a full internal audit and three robustness tests. A number of analyses still remain, and several points require correction or carry standing caveats. They are listed here in full, so that no result is relied upon beyond what the present evidence supports, grouped into remaining analyses, corrections to be made before publication, and standing limitations.

Remaining analyses. The rolling-origin forward test has been performed on HS2 itself (Section 5.7) and the multi-seed replication stands (Section 6.10); what neither can establish, because HS2's outcome is known, is prospective accuracy, and that requires running the same discipline, under the same cut-off rule, on a programme whose outcome is not yet settled.

Standing limitations. Total programme slip is independently sourced; the allocation of that slip to individual schedule systems is an informed model (Section 4.1), so the per-system day figures should be read as modelled, though the structural gap survives this, depending on ordering rather than absolute calibration. Three cost systems, utilities (HC03), stations (HC05) and railway systems (HC06), have no separately published Phase 1 magnitude and carry conservative placeholders, which makes the modelled cost an understatement on those systems. The 13.6 per cent duration-sensitive share in the schedule-to-cost coupling (Table 4) is the single load-bearing modelling assumption; it is bounded below by a sourced floor but is not itself separately published, and the coupling's magnitude scales with it. And the reference-class parameters (Flyvbjerg, Holm and Buhl, 2002) are two decades old and should be refreshed against more recent syntheses; more fundamentally, the method is demonstrated here on a single programme whose outcome was already known, so it is a worked case establishing mechanism and plausibility, not a validated forecasting instrument, and it should not be cited as the latter.

8. Conclusion

The subject of this paper is a method, and the conclusion is about the method. The argument was that two inexpensive modelling practices (letting a risk register express the dependence between its risks, and letting the baseline learn as evidence arrives) recover risk that the conventional approach suppresses, and that the amount recovered can be measured. HS2 served as the worked case because its registers are public and its outcome is known, which together allow the method to be checked against a target. The case should not be read as the result; the method is the result.

On that method, three findings stand. First, the independence assumption has a price that can be measured on a project's own register, by running it twice: on HS2 the price was about £21.3bn of cost exposure and about 864 days of schedule exposure, each stable under resampling and each independent of any forecast made after the fact, because each is the difference between two runs on identical magnitudes. Second, coupling the schedule and cost models through a single documented node makes the conversion of programme slip into money explicit and auditable, and on HS2 that node (derived, not assumed) turned out to be one of the larger contributors to the cost answer. Third, when the dependence is admitted, the model reaches the very band that the programme's own reset later confirmed (the three dependency-aware schedule lenses all falling inside the ten-to-seventeen-year slip, and the cost chain reaching the floor of the £87.7–102.7bn range from the coupled register alone) which indicates that the reset was implicit in the project's registers from the start, waiting only for the dependence between the risks to be switched on. That updating the chain on evidence then moves it only modestly confirms that the structure, not the later data, set the level.

The method's appeal is that it is cheap and general. It needs no bespoke software and no data a project does not already hold; the marginal cost is one additional simulation. Any sponsor maintaining a quantitative risk register can compute its own independence error, in its own currency and its own units of time, and can report that figure as a measure of how much risk its current model is concealing. A project that cannot state its independence error does not yet know its own exposure.

One caveat closes the argument honestly. The method is demonstrated here on a single programme with a known outcome, so it establishes mechanism and plausibility rather than validated predictive accuracy, and it should be cited as the former. The rolling-origin forward test has now been run on that programme (Section 5.7), showing what the method would have reported as each document was published: the structure reproduced at 2013, the independence signature at 2020, and the climb into and past the reset band as magnitudes were disclosed. What it cannot show, because the outcome is already known, is prospective accuracy; the decisive remaining step is to run the same discipline forward, under strict information cut-offs, on a programme whose outcome is not yet settled. What this paper establishes is narrower but, I hope, secure: that the independence assumption is not free, that its price is measurable on a project's own register, that the price survives every reasonable change of graph and distribution, and that on HS2 it was large; and that the register, run twice, would have said so at the outset. The register has no memory. The modeler must supply it.

Author's note on data used: All source data is drawn from public UK Government, National Audit Office and Parliamentary risk documents; no company data is used. The runs are reproducible at seed 4209 on the RiskPulse engine (build V12.5.2), at 100,000 iterations and seed 4209 on every run. Every figure quoted was recomputed from the raw simulation output by a separate check, independently of the engine that produced it, and all such checks pass. The full configuration (the probability and three-point range on every system, and the conditional probability on every arrow, for both the schedule and the cost network) is given in Appendix C. The observation registers, the topology variants and the per-node breach output are released as workbooks alongside this paper.

References

References follow the author - date convention. Each is cited at a specific claim in the text; no general-reading entries are included. Government and audit sources are identified by their official document numbers; the principal cost and schedule figures are drawn from these primary sources.

Committee of Public Accounts (2025) *High Speed 2.* HC 357. London: House of Commons. Available at: www.parliament.uk.

Department for Transport (2023) *High Speed 2: six-monthly report to Parliament.* Written ministerial statement HCWS34, November 2023. London: UK Parliament. Available at: www.gov.uk.

Department for Transport (2025) *High Speed 2: six-monthly report to Parliament.* Written ministerial statement HCWS852, July 2025. London: UK Parliament. Available at: www.gov.uk.

Department for Transport and HS2 Ltd (2013) *HS2 Cost and Risk Status Report (Schedule and Accounts 8).* October 2013. Available at: www.gov.uk. [Independence assumption: clause 4.2.3; component magnitudes: Table 16, p. The choice was orthodox practice for its period, and the report states it with unusual candour; this paper prices the assumption, it does not prosecute it. 43.]

Department for Transport and HS2 Ltd (2020) *Full Business Case: High Speed 2 Phase One.* April 2020. Available at: www.gov.uk.

Dhand, N. (2026) 'The £100 Billion Blind Spot', *PM World Journal*, Vol. XV, Issue VI, June 2026. https://pmworldlibrary.net/wp-content/uploads/2026/06/pmwj165-Jun2026-Dhand-The-100-Billion-Blind-Spot.pdf

Embrechts, P., McNeil, A.J. and Straumann, D. (2002) 'Correlation and dependence in risk management: properties and pitfalls', in Dempster, M.A.H. (ed.) *Risk Management: Value at Risk and Beyond.* Cambridge: Cambridge University Press, pp. 176–223.

Fenton, N. and Neil, M. (2018) *Risk Assessment and Decision Analysis with Bayesian Networks.* 2nd edn. Boca Raton: CRC Press. [Noisy-OR parameterisation; Bayesian-network risk modelling.]

Flyvbjerg, B. and Gardner, D. (2023) *How Big Things Get Done.* London: Macmillan.

Flyvbjerg, B., Holm, M.S. and Buhl, S. (2002) 'Underestimating costs in public works projects: error or lie?', *Journal of the American Planning Association,* 68(3), pp. 279–295. [Rail reference class: mean overrun ≈ 44.7%, SD ≈ 38.4%.]

HS2 Ltd (2012) *HS2 Cost and Risk Model Report.* Available at: www.gov.uk. [System risk descriptions: Appendices B and D.]

HS2 Ltd and Department for Transport (2026) *Accounting Officer Assessment: High Speed 2.* May 2026. Available at: www.gov.uk. [2026 reset: cost band £87.7–102.7bn; opening 2036–2043; spend to March 2026 £44.2bn.]

Iman, R.L. and Conover, W.J. (1982) 'A distribution-free approach to inducing rank correlation among input variables', *Communications in Statistics – Simulation and Computation,* 11(3), pp. 311–334.

ITV News (2026) *HS2 spiraling costs: slower trains, less than half the size.* 19 May 2026. Available at: www.itv.com . [De-specification: design speed reduced from 360 to 320 km/h; automatic train operation dropped.]

Jensen, F.V. and Nielsen, T.D. (2007) *Bayesian Networks and Decision Graphs.* 2nd edn. New York: Springer.

Kahneman, D. (2011) *Thinking, Fast and Slow.* New York: Farrar, Straus and Giroux. [Optimism bias and the planning fallacy.]

Müller, A. and Stoyan, D. (2002) *Comparison Methods for Stochastic Models and Risks.* Chichester: Wiley. [Comparison of distributions at matching percentiles.]

National Audit Office (2016) *Progress with preparations for High Speed 2.* HC 235. London: NAO. Available at: www.nao.org.uk.

National Audit Office (2020) *High Speed 2: a progress update.* HC 40. London: NAO. Available at: www.nao.org.uk.

National Audit Office (2024) *High Speed 2: update following cancellation of Phase 2.* HC 128. London: NAO. Available at: www.nao.org.uk.

New Civil Engineer (2026) *Report on the HS2 de-specification.* 23 May 2026. [Corroborates the ITV disclosure entered as a saving on HC10.]

Pearl, J. (1988) *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* San Francisco: Morgan Kaufmann.

Rockafellar, R.T. and Uryasev, S. (2000) 'Optimization of conditional value-at-risk', *Journal of Risk,* 2(3), pp. 21–41. [Definition of CVaR / tail mean.]

Vose, D. (2008) *Risk Analysis: A Quantitative Guide.* 3rd edn. Chichester: Wiley. [PERT sampling; correlated Monte Carlo practice.]

Appendix A: Glossary

Bayesian network. A set of uncertain quantities (nodes) joined by directed arrows, where an arrow encodes that one quantity bears causally on another. Used here to represent the dependence between project risks explicitly, rather than through a correlation coefficient applied afterwards.

Noisy-OR. A rule for combining several causes acting on one effect: each firing cause makes an independent attempt to trigger the effect, which fails to occur only if every firing cause independently fails. It produces diminishing returns as causes accumulate and never reaches certainty.

PERT (beta-PERT) distribution. A continuous distribution specified by a minimum, most-likely and maximum value, widely used for three-point cost and duration estimates. The shape parameter $\lambda = 4$ is the conventional default.

Iman–Conover method. A distribution-free technique for inducing a target rank correlation between simulated inputs. Here it is switched on but targets a zero matrix, so it imposes no residual correlation and all dependence is structural.

P50 / P80 / P90. Percentiles of the output distribution: the value below which 50, 80 or 90 per cent of simulated outcomes fall. P90 is the high-confidence figure reported throughout.

CVaR(90%) (conditional value at risk; tail mean). The average of all outcomes worse than the P90. It measures the severity of the tail, not only its threshold.

Independence error. The difference between the dependency-aware P90 and the independent P90 on identical magnitudes; the amount of risk a model suppresses by assuming its risks are independent. The central measured quantity of this paper.

Convergence node. A node on which several arrows land. Once a network is active, convergence nodes become near-certain to fire, which both concentrates risk and relocates management purchase onto the arrows feeding them.

Coupling node (HC12). The single cost node whose magnitude is derived from the schedule network's programme slip, by which schedule risk enters the cost answer through a documented mechanism.

Appendix B: Declaration on data, prior work, and the use of AI

Data. Every figure in this paper derives from public sources: UK Government departmental reports and written statements, National Audit Office reports, and Parliamentary committee reports, identified individually in the References and in the observation registers released alongside this paper. No commercial, confidential or personal data is used. The principal simulation runs are reproducible at random seed 4209 and 100,000 Latin-hypercube iterations; each configuration carries an input digest that fixes its register exactly, and every headline figure is reconstructed from the raw output in Section 6, independently of the simulation engine.

Data availability. Every quantitative claim in this paper is reproducible from materials lodged with the submission: the two Bayesian network registers, schedule and cost, each carrying its node table, its dependency table and its sources sheet; the four rolling-origin intelligence logs, cost official, cost modelled, schedule official and schedule modelled, each carrying its run provenance, its per cut-off configuration hash and its full observation diary; the five-seed multi-seed workbook; the companion source provenance dossier, which sets every register input beside the document and locator it derives from; and the engine build, RiskPulse V12.5.2. Seed 4209 at 100,000 iterations reproduces every headline figure, and Appendix C carries the full configuration.

Relationship to prior work. The clause 4.2.3 argument was first put to this journal's readers by the present author in Dhand (2026), which illustrated it on a seven-node, author-estimated table using an earlier build of the same engine, RiskPulse. That paper and this one therefore share a foundational premise and a tool lineage, and both are declared here. This paper is not a restatement of it. The registers are rebuilt to Phase 1 from the primary 2013 source, the schedule domain and the schedule-to-cost coupling are new, and the forward test and the four robustness tests of Section 6 are new. The engine is the author's own work, and the version increment to build V12.5.2, together with the defects corrected in it, is documented in Section 6.2.

Use of AI. Generative AI tools were used for the structural organisation and language editing of this manuscript only: arranging sections into the conventional academic order, tabulating figures already computed. All modelling, analysis, data, figures, results and conclusions are the author’s own. The author has reviewed and verified every statement and accepts full responsibility for the accuracy of the analysis, the validity of the sources, and the final wording.

Appendix C: Full network configuration

C.1 Engine configuration and node magnitudes

Both networks were evaluated on the RiskPulse engine (see Section 6.2 on version labelling) with the constants common to every run: random seed 4209; 100,000 iterations; Latin-hypercube sampling; PERT shape λ = 4; noisy-OR enabled on all networked runs; and the residual-correlation stage enabled but targeting a zero matrix. The node magnitudes (the marginal firing probability and three-point range for each of the ten schedule systems and the twelve cost systems) are given in full in Tables 2 and 3, and the derivation of the coupling node HC12 in Table 4. The dependency structure, summarised by figure in the body, is specified completely below: each arrow carries the conditional probability that the child fires when its parent has fired, and when its parent has not.

One configuration constant deserves its own line. The Iman–Conover correlation stage is present in every dependency-aware run with its target matrix set to the zero matrix; every off-diagonal target is 0.00. The stage is left switched on so that its sampling bookkeeping is identical across configurations, and its measured contribution to the P90 is about £0.003bn (Section 6.4). Dependence in this paper is carried by the arrows, not by induced rank correlation.

C.2 Schedule dependency structure (13 arrows)

From	To	P(child fires | parent fires)	P(child fires | parent safe)
HS10 Funding	HS07 Contractor	0.80	0.20
HS10 Funding	HS01 Land	0.75	0.20
HS10 Funding	HS08 Scope	0.70	0.25
HS01 Land	HS07 Contractor	0.82	0.18
HS01 Land	HS03 Utilities	0.70	0.25
HS06 Environmental	HS01 Land	0.78	0.20
HS06 Environmental	HS07 Contractor	0.72	0.20
HS03 Utilities	HS07 Contractor	0.72	0.22
HS02 Ground	HS04 Main civils	0.80	0.18
HS09 Governance	HS08 Scope	0.75	0.25
HS09 Governance	HS10 Funding	0.70	0.25
HS08 Scope	HS07 Contractor	0.70	0.25
HS04 Main civils	HS05 Stations	0.75	0.20

Table C.1. The thirteen dependency arrows of the schedule network. *Five arrows converge on contractor performance (HS07).*

C.3 Cost dependency structure (15 arrows)

From	To	P(child fires | parent fires)	P(child fires | parent safe)
HC01 Land	HC03 Utilities	0.85	0.15
HC02 Ground	HC04 Main civils	0.80	0.15
HC03 Utilities	HC04 Main civils	0.78	0.15
HC09 Contractor	HC04 Main civils	0.82	0.15
HC10 Scope	HC04 Main civils	0.85	0.15
HC04 Main civils	HC05 Stations	0.80	0.20
HC10 Scope	HC05 Stations	0.82	0.20
HC04 Main civils	HC12 Schedule cost	0.85	0.18
HC07 Environmental	HC12 Schedule cost	0.78	0.18
HC11 Governance	HC12 Schedule cost	0.75	0.18
HC12 Schedule cost	HC08 Inflation	0.88	0.50
HC11 Governance	HC10 Scope	0.82	0.20
HC11 Governance	HC09 Contractor	0.80	0.20
HC05 Stations	HC06 Railway systems	0.78	0.18
HC08 Inflation	HC06 Railway systems	0.75	0.18

Table C.2. The fifteen dependency arrows of the cost network. *Four arrows converge on main civils (HC04); the coupling node HC12 sits downstream of main civils and feeds inflation (HC08).*

C.4 A note on distributional shape

A feature of the outputs worth a word is that the dependency-aware distributions are close to symmetric, and slightly left-skewed, where intuition might expect a long right tail. The schedule chain has a skewness of about +0.04 and the cost chain about −0.05, against roughly +0.1 to +0.2 for the corresponding independent runs. The mechanism is the one described in Sections 5.1 and 5.3: under dependence, the downstream convergence nodes fire on almost every iteration, so the outcomes that would have populated the upper tail of the independent distribution instead become typical, pulling the body of the distribution up toward the former tail and compressing the skew. This is why the median rises more than the P90 between the independent and dependency-aware lenses, and why the P90-to-P50 ratio falls from about 1.46 to about 1.20 on cost. It is a property to expect, not an anomaly to correct: a strongly coupled system concentrates its outcomes, because its risks arrive together rather than in isolation.

Appendix D: The component-breach test in full

The test of Section 6.9 reads, for each driver system, its own independent P90 ceiling from the independent-by-node run, and then counts how often the modelled-by-node run exceeds that ceiling; system by system and several at once. Under independence each driver breaches its own ceiling ten per cent of the time by construction; the modelled column shows the lift the dependence produces. Tables D.1 and D.2 give the full per-driver detail for the cost and schedule networks.

Cost driver	Indep. P90 ceiling (£bn)	Breach rate, indep.	Breach rate, modelled	Lift
Main civils	22.9	10.0%	14.9%	1.5×
Schedule delay cost (HC12)	17.6	10.0%	17.4%	1.7×
Stations and depots	10.5	10.0%	20.4%	2.0×
Inflation and funding	14.7	10.0%	18.0%	1.8×
Scope and design	10.1	10.0%	13.1%	1.3×
Contractor and supply chain	9.8	10.0%	12.0%	1.2×

Table D.1. Cost network. *Among these six drivers, the expected number of simultaneous breaches rises from 0.600 to 0.958 and its variance from 0.539 to 0.804; the all-six joint breach is about one in fifty thousand. The same test against the pure chain gives a larger lift (mean 1.03, variance 0.85).*

Schedule driver	Indep. P90 ceiling (days)	Breach rate, indep.	Breach rate, modelled	Lift
Contractor performance	583	10.0%	21.2%	2.1×
Scope and design	408	10.0%	15.9%	1.6×
Stations (Euston)	945	10.0%	14.5%	1.5×
Main civils	886	10.0%	11.1%	1.1×
Funding instability	856	10.0%	10.9%	1.1×
Governance and reset	601	10.0%	8.1%	0.8×

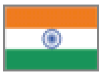
Table D.2. Schedule network. *The expected number of simultaneous breaches rises from 0.600 to 0.817 and its variance from 0.538 to 0.697. Every driver lifts except governance, whose prior the evidence revised downward (Section 5.4), so that it breaches slightly less than its independent ceiling implies; reported as found.*

The structural-robustness sweep, the distributional test and the component-breach test reported above complete the validation set the reviewers requested.

About the Author

Nikhil Dhand

Delhi, India



Nikhil Dhand, PMP is a Senior Project Controls Consultant with 11 years of cross-sector experience delivering projects for UK, US, and Indian clients across solar EPC, highways, urban infrastructure, and complex project delivery. He is the creator of Probabilistic Chain Analysis™ (PCA) and RiskPulse V12™ — a Bayesian Monte Carlo risk engine calibrated using data from 16,000+ infrastructure projects across 8 sectors.

Author of *Risk Runs the Project (forthcoming, Amazon KDP).*

ORCID: 0009-0000-6526-8399 | **Zenodo:** 10.5281/zenodo.19436269

Contact: nikhildhand.research@gmail.com |