AI Engineering ROI Calculator — Model Methodology and Audit Specification

Version 1 | Controlled draft for independent methodological audit

31 October 2026

Table of Contents

Document control 3
Executive summary 3
1. Objective, population, counterfactual and boundaries 4
1.1 Decision question 4
1.2 Counterfactual (BAU) 4
1.3 Reporting scope 4
2. Consistent notation and calculation units 4
3. Workforce, currency and full employment costs 6
3.1 Workforce cohorts 6
3.2 Cost conventions 7
3.3 Illustrative cost defaults — not audited benchmarks 7
3.4 Exchange rates 7
4. Incremental AI cost model 7
5. AI productivity and adoption 8
5.1 Two exclusive productivity-input modes 8
5.2 Adoption 8
5.3 Redeployment of released capacity 9
6. Separate positive and negative productivity channels 9
7. Capacity-based economic value 9
8. Revenue-linked contribution-value model 10
8.1 Company financial baseline 10
8.2 Commercial response coefficient 10
8.3 Commercial diffusion loss 11
8.4 Cohort aggregation 11
8.5 Revenue and contribution calculations 11

8.6 Profit and attribution discipline ... 12
9. Net benefit and return measures ... 12
10. Full defaults and mandatory classifications ... 12
10.1 Baseline fictional company ... 12
10.2 Industry presets — unvalidated demonstration values ... 14
10.3 Data provenance requirements ... 14
11. Numerical reference case — independently reproducible ... 14
11.1 Single-variable productivity sensitivity ... 15
11.2 Valuation reversal threshold ... 16
11.3 Negative stress case ... 16
12. Calculation state, UI, saving and exports ... 16
12.1 User flow ... 16
12.2 Result hierarchy ... 17
12.3 Versioned immutable snapshots ... 17
12.4 Sharing and confidentiality ... 17
13. Validation, invariants and acceptance tests ... 17
13.1 Input validation ... 17
13.2 Required model invariant tests ... 18
13.3 Property-based tests ... 19
13.4 Auditability ... 19
14. Known limitations and model risk register ... 20
15. Evidential standards and calibration plan ... 22
15.1 Priority evidence for a real organisation ... 22
15.2 Research context (not validation of defaults) ... 23
15.3 Calibrating the industry library ... 23
16. Scope decisions and change log ... 23
16.1 In scope — financial engine v1 ... 23
16.2 Deferred until explicitly re-specified ... 23
16.3 Changes from v1.0 to v1 ... 24
17. Independent audit protocol and acceptance gate ... 24
18. Primary references and interpretation ... 25

Document control

Field	Value
Document ID	DODONE-MOD-001
Version	1
Status	**For independent review**
Scope	One-year, incremental ROI of deploying AI-assisted software-engineering tools
Reporting basis	Company economic value and incremental tool expenditure, not necessarily audited cash flow
Intended implementation	Deterministic TypeScript engine; web application independent of methodology
Last Review date	31 October 2026

Executive summary

The calculator estimates the **incremental economic effects** of using an AI coding tool compared with a defined **business-as-usual (BAU) scenario without that tool**. The analysis includes incremental licence fees, separately charged usage/tokens and optional deployment costs. It evaluates productivity through the share of engineering work affected, net time saved per affected unit of work, adoption, and the proportion of the released capacity **actually redeployed into comparable additional engineering output**.

The same realised output is valued by **two alternative, mutually exclusive methods**:

1. **Capacity-based economic value**: the cost-equivalent value of released engineer time that is redeployed into productive additional work, using fully loaded compensation.
2. **Revenue-linked contribution value**: an explicitly hypothetical commercial outcome using counterfactual annual revenue, a commercial responsiveness factor for engineering, additional realised engineering output, a diffusion loss for downstream commercial constraints, and an incremental contribution margin.

These methods are **not additive, not guaranteed lower/upper bounds, and not equivalent to booked cash savings or independently established incremental profits**. Both result in distinct net benefit and ROI estimates after the same incremental cost is deducted. Under the specified default fictional company, annual costs are **£31,800**, capacity benefit **£253,500**, and commercial contribution benefit **£609,375**; the respective indicative ROI estimates are **697.2%** and **1,816.3%**.

The biggest limitation is that the commercial method is a **scenario-valued proxy** for the marginal relationship between engineering output and sales, not a causal revenue forecast. The proposed industry and location defaults are demonstration values pending external validation.

1. Objective, population, counterfactual and boundaries

1.1 Decision question

For a specified company and a single 12-month analysis period, **what is the incremental annual economic value, under transparent assumptions, of deploying an AI coding tool to an identified engineering workforce, compared with the same workforce and operations without that incremental AI deployment?**

The model does not calculate a company's total return on AI, general AI transformation, organisational value or long-term enterprise valuation.

1.2 Counterfactual (BAU)

For v1 the BAU case assumes the *incrementally assessed tool is not in use* and includes the engineering team's existing output, expected ordinary revenue and margins, and existing compensation and operations. BAU productivity improvements unrelated to the assessed AI tool are **excluded from the AI benefit**. The user supplies last financial year's revenue and a baseline growth assumption to forecast BAU revenue for the analysis year; default growth is 0%.

Important eligibility restriction: the reference formulas directly apply to a new adoption versus a no-tool baseline. Organisations replacing an existing AI tool, changing supplier or increasing an existing AI rollout cannot use the formulas without separately defining **incremental adoption, incremental task productivity and incremental costs relative to the incumbent BAU**. This is a future enhancement, not silently assumed.

1.3 Reporting scope

The model presents an annual **economic-value appraisal**, not a conventional cash-flow investment case. Fully loaded salary measures the opportunity cost of engineering effort; existing salaries are not incremental AI expenditure. Thus capacity benefit is not booked payroll saving. Revenue-linked benefit is a contingent *incremental contribution-profit opportunity*. Neither should be reported as audited realisation without observed evidence.

Year-one average FTE counts, AI adoption and costs are used throughout. In-year ramp profiles and recurring revenue persistence are explicitly outside v1.

2. Consistent notation and calculation units

All percentages are stored as decimal proportions, not percentage numbers. Monetary values refer to the same company, currency and appraisal period. Engineering counts are annual-average full-time equivalents (FTE), which may be fractional. All productivity measurements are **net of time spent prompting, reviewing, testing, reworking, securing and integrating AI-assisted output**, so the model does not add an overlapping quality penalty.

Symbol	Definition	Unit / bounds
N	Total baseline engineering workforce	Annual-average FTE; $N \geq 0$
n_i	Engineers in workforce cohort i (location × seniority)	FTE; $n_i \geq 0$; $\sum_i n_i = N$
w_i	Fully loaded annual cost per engineer in cohort i	Reporting currency / FTE-year; $w_i \geq 0$
a_i	Fraction of engineers in cohort i effectively active with AI	[0,1]; constrained by licence coverage
f_i	Fraction of baseline engineering work time eligible for AI	[0,1]
s_i	**Net time reduction** on AI-eligible work, relative to BAU time	$[-0.5,0.8]$ for v1; negative means slower
r_i	Fraction of total baseline time released by an active user	$r_i = f_i s_i$, so $-0.5 \leq r_i < 1$
g_i	Equivalent end-to-end completed-output uplift for active user at constant time budget	$g_i = r_i/(1 - r_i)$; $g_i > -1$
u_i	Share of potential additional engineering production **actually undertaken** using released time	[0,1], applied to positive productivity only
R_0	Company revenue in most recently completed fiscal year	Reporting currency / year; $R_0 \geq 0$
b	BAU annual revenue growth rate	Decimal; v1 default 0
R	Forecast revenue **without the assessed AI** for analysis year	$R = R_0(1 + b)$; $R \geq 0$
e	Engineering commercial-response proxy / addressable revenue fraction	[0,1] in simplified v1

Symbol	Definition	Unit / bounds
d	Positive-output downstream commercial **diffusion loss**	[0,1]; $1 - d$ retained
t_-	Downside commercial transmission if AI reduces engineering output	[0,1]; default 1
m	Incremental contribution margin on attributable additional sales	$[-1,1]$; default from industry illustrative proxy
L	Average paid AI licences	Number of annual-average seats; $0 \leq L \leq N$
p_L	Price per licensed seat per month	Reporting currency; $p_L \geq 0$
T	Monthly additional usage charge per active AI engineer	Reporting currency; $T \geq 0$
I	Additional year-one rollout / training / integration / governance expenditure	Reporting currency; $I \geq 0$
C	Total incremental year-one AI deployment cost	Reporting currency

Use of e: the public interface may call it *Engineering commercial contribution*. For financial accuracy it is defined as the **fraction of BAU annual revenue assumed to respond proportionally to marginal engineering output**, before downstream loss. It is **not an observed accounting allocation** of historic gross profit. When historic and incremental margin are identical, the equivalent presentation $R \times m \times e$ may be called *engineering-addressable contribution pool*, but this naming does not establish causal attribution.

3. Workforce, currency and full employment costs

3.1 Workforce cohorts

The authoritative input is a matrix of annual-average engineering FTEs by location and seniority. Summarised location and seniority percentages may be used to create that matrix for novice users. **Do not assume location and seniority are statistically independent** if the user specifies a detailed matrix; detailed cohort counts always take precedence. Calculate:

$$N = \sum_i n_i\,, \qquad W = \sum_i n_i\, w_i, \qquad \bar{w} = \begin{cases} W/N, & N > 0 \\ 0, & N = 0. \end{cases}$$

Here W is baseline annual loaded engineering payroll and $\bar{w}$ its weighted unit cost. Existing payroll is never counted as an incremental AI tool cost.

3.2 Cost conventions

Loaded employment costs should consistently include salary, expected incentive compensation, employer taxes, employee benefits and the configured allocation of direct employment overheads. Every reference rate records which components it includes. Avoid including an overhead amount twice through both a loaded salary and a separate expense. Compensation may be entered directly as a fully loaded value or derived from salary and a documented loading percentage.

3.3 Illustrative cost defaults — not audited benchmarks

Engineer location	Junior	Mid-level	Senior
United Kingdom	£52,000	£78,000	£117,000
United States	$124,000	$182,000	$247,000
Poland	PLN 120,000	PLN 225,000	PLN 325,000
India	₹875,000	₹2,250,000	₹3,750,000

Reference location: United Kingdom. Reference seniority mix: 20% junior / 50% mid / 30% senior. Reference location mix: 100% UK. An *Other / Custom* location must be available, with a manual currency and cost entry.

These figures are **artificial starting values used for product testing**. They must be marked *illustrative — source not validated*, and must not be described as 2026 compensation survey statistics.

3.4 Exchange rates

Convert individual cohort salaries into the **reporting currency at stored, dated rates before aggregation**. Store original amount, original currency, conversion rate, rate source/date and any manual override. FX differences and purchase-power effects are not modelled within the 12-month horizon. Conversion must never silently modify the original local compensation assumption.

4. Incremental AI cost model

For one annual-average period:

$$C_L = 12Lp_L$$

$$C_T = 12\left(\sum_i n_i\, a_i\right)T$$

$$C = C_L + C_T + I$$

T represents additional out-of-plan spend after any included credits, not total tokens generated or provider cost of service. If per-cohort usage prices vary, substitute a cohort-specific T_i and sum accordingly. Subscription seats and active engineers are distinct concepts; generally the number of active engineers must not exceed licensed engineers for a tool requiring paid seats. In simple mode, use totals; in cohort mode, validate eligibility at cohort level.

For a 12-month calculation, multiplying annual-average seats by 12 assumes seats and consumption are spread evenly. It cannot estimate a detailed ramp or annual commitment true-up. Taxes, exchange-rate effects on vendor billing, infrastructure, governance and implementation are excluded unless explicitly included in the relevant cost field. The default $I = 0$ **is not a factual assertion that deployment has no cost**.

When $C = 0$, compute and show benefits and net value, but report ROI **not defined**. Do not report an infinite ROI. When net benefit is negative, ROI can be below −100%; do not artificially floor it.

5. AI productivity and adoption

5.1 Two exclusive productivity-input modes

Mode A — task time saved. Enter AI-eligible fraction of baseline work f_i and the *net* fractional time reduction on eligible work s_i. Then:

$$r_i = f_i s_i, \qquad g_i = \frac{r_i}{1 - r_i}.$$

This assumes the **same mix and quality of comparable engineering work**, a stable baseline time requirement and no hidden task overheads. For example, if 50% of engineering time is eligible and eligible work takes 20% less time, $r_i = 10\%$ and maximum comparable output uplift at constant labour hours is $g_i = 11.111 \ldots \%$. A **20% task time reduction is not a 20% entire-team output uplift**.

Mode B — observed overall output uplift. Enter measured, quality-adjusted end-to-end output uplift g_i, relative to the same BAU engineer/time unit. Then:

$$r_i = \frac{g_i}{1 + g_i}.$$

The app must **not multiply Mode B output uplift by the eligible work fraction a second time**. Negative changes are permitted; $g_i > -1$ and the implied r_i must remain in the supported time-saving bounds.

5.2 Adoption

a_i is the fraction of the cohort that consistently achieves the stated AI productivity change **during the annual-average analysis period**. Licensing is neither evidence of use nor

evidence of a productivity benefit. Defaults: annual-average 60% active, 70% licensed. Subject to product licence eligibility, $0 \leq a_i \leq 1$.

5.3 Redeployment of released capacity

u_i is **not** the adoption rate or the success of sales. It is the fraction of the *potential additional engineering work* that is actually produced by reinvesting the time released through AI. The simplifying micro-assumption is that the same fractional amount of theoretically freed first-pass time is redeployed into additional comparable tasks at the new AI-assisted productivity rate. Under that assumption the realised output uplift for positive r_i is $a_i u_i g_i$ at cohort level.

This definition is deliberate: $a_i u_i g_i$ is an approximate realised-output measurement, **not** $a_i r_i u_i/(1 - r_i u_i)$. The latter corresponds to a different time-reinvestment mechanism and must not be substituted without changing the model definition. The default $u_i = 50\%$ is illustrative and is not a directly measured national or industry statistic.

If a user reallocates saved time to training, risk reduction or quality improvements, its monetary benefit is **not automatically captured** unless it is expressed as comparable additional engineering output. Cash savings from genuinely avoided hiring or payroll are also a separate potential benefit channel; v1 does not add them to these two valuations.

6. Separate positive and negative productivity channels

The model must not allow a *positive-benefit discount* to conceal the consequences of an adverse productivity outcome. For each cohort define:

$$r_i^+ = \max(r_i, 0), \qquad r_i^- = \min(r_i, 0)$$

$$g_i^+ = \max(g_i, 0), \qquad g_i^- = \min(g_i, 0).$$

For positive output, apply the cohort's realised production factor u_i. For negative output, assume the adverse performance effect is fully incurred among the active users unless a separate downside avoidance or mitigation case is explicitly modelled. Commercial negative transmission t_- is a **different** parameter from positive diffusion loss d; default $t_- = 100\%$.

This is a conservative downside convention, not an empirical finding. If a company can immediately stop using a failing tool, the true downside may be smaller; that requires a distinct intervention/cessation assumption. The negative method is a scenario stress case, not a claim of observed realised losses.

7. Capacity-based economic value

For cohort i:

$$B_{C,i} = n_i w_i a_i (u_i r_i^+ + r_i^-).$$

Total capacity-value benefit:

$$\boxed{B_C = \sum_i B_{C,i}}$$

For positive productivity, this measures the **compensation-cost equivalent of the released baseline labour time that is used to perform additional useful engineering work**, not a payroll expenditure reduction. For negative productivity, it reflects the estimated cost equivalent of time lost under the downside convention.

If $N = 0$, $B_C = 0$. If $a_i = 0$ or $r_i = 0$ for all engineers, $B_C = 0$. B_C can be negative. This result is never described as a cash saving without separate evidence of avoided actual expenditure. It is not an accounting profit and should not be added to commercial benefit.

8. Revenue-linked contribution-value model

8.1 Company financial baseline

R_0 is the last completed financial year's revenue. For the 12-month appraisal year, define BAU revenue:

$$R = R_0(1 + b).$$

b is the **ordinary growth expected without the assessed incremental AI deployment**. Default $b = 0\%$. Historical gross margin can be displayed as a useful reference, but the model's marginal conversion from revenue to economic contribution uses **incremental contribution margin** m, which includes all variable sales/delivery costs associated with incremental revenue. If m is not known, the illustrative industry margin is a clearly labelled proxy.

8.2 Commercial response coefficient

e describes the share of BAU revenue *assumed to be sensitive to additional engineering output* at the margin, before diffusion loss. It is constrained to [0,100%] in v1 for interpretability. This bound is a design simplification, not a theorem about business elasticities.

Using historic accounting profit divided by the number of engineers is a helpful descriptive diagnostic but does **not** establish the marginal effect of additional engineering work on sales. Therefore **never label e as a factually measured share of gross profit caused by engineering** unless a separate causal attribution study justifies it.

The UI can display the diagnostic:

$$\text{Engineering addressable contribution per FTE} = \frac{R \times e \times m}{N}$$

when $N > 0$. This is a *model-derived reference value* and not an audited financial reporting allocation.

8.3 Commercial diffusion loss

d is the proportion of potential *positive incremental sales opportunity* lost within the appraisal period due to customer demand, release cadence, fulfilment, sales, regulatory, platform or other **downstream non-engineering constraints**. It does **not** repeat AI non-adoption, failure to use released engineer time, or rework already captured in net productivity. Default is 25% for the illustrative SaaS industry.

Because the model contains only one year, d **treats all unrealised current-year sales as not achieved in that year**, including sales that may instead be deferred. Multi-year recovery/persistence is excluded. The factor is therefore a same-period conversion parameter, not a claim that all lost value disappears forever.

8.4 Cohort aggregation

Define average *realised positive* and *adverse* engineering-output uplift across all engineers:

$$Q_+ = \begin{cases} \frac{1}{N}\sum_i n_i\, a_i u_i g_i^+, & N > 0 \\ 0, & N = 0 \end{cases}$$

$$Q_- = \begin{cases} \frac{1}{N}\sum_i n_i\, a_i g_i^-, & N > 0 \\ 0, & N = 0. \end{cases}$$

$Q_+ \geq 0$ and $Q_- \leq 0$. In the simplest case of uniform positive productivity, $Q_+ = aug$.

8.5 Revenue and contribution calculations

$$\boxed{\Delta R = R\, e[Q_+(1-d) + Q_- t_-]}$$

$$\boxed{B_R = \Delta R \times m}$$

ΔR is a **scenario-modelled change in revenue**, potentially negative; B_R is the **scenario-modelled incremental contribution profit**. A negative incremental contribution margin can make additional sales economically detrimental, so the input's sign must remain visible. Do not force B_R to be positive or claim this is a reliable statistical prediction.

In the default positive-productivity case:

$$\boxed{B_R = R \times m \times e \times a \times \frac{r}{1-r} \times u \times (1-d).}$$

The product of $e(1-d)$ is the **effective commercial response coefficient**. Because only their product is identified by the final financial result, the two factors must not be

portrayed as independently estimated from the output. They require separate evidence or remain explicit judgemental assumptions.

8.6 Profit and attribution discipline

- **Never** apply both historical gross margin and incremental contribution margin to the same additional sales.
- **Never** treat commercial gross-profit / contribution-profit opportunity as identical to incremental revenue.
- **Never** add B_R to B_C or imply that the higher one is a proven upper bound.
- The company must have meaningful monetisable engineering output; if there is no viable additional demand, set $e = 0$ or $d = 100\%$ on the positive channel.
- Commercial output may depend on price, market share, product-market fit, launch timing, execution quality and customer behaviour; these are simplified into e and d, not explicitly forecast.

9. Net benefit and return measures

With total incremental cost C:

$$V_C = B_C - C, \qquad V_R = B_R - C$$

When $C > 0$:

$$\boxed{ROI_C = 100\% \times \frac{B_C - C}{C}}$$

$$\boxed{ROI_R = 100\% \times \frac{B_R - C}{C}}$$

Also disclose benefit-to-cost ratios B_C/C and B_R/C when meaningful. ROI describes **a ratio of scenario-derived economic value to incremental expenditure**, not necessarily a conventional audited financial ROI. If $C = 0$, ROI is "not defined" and benefits / net value are still displayed. If a negative benefit exceeds cost, ROI may be less than −100%.

Display the valuations side by side, **not summed**, and do not mechanically label capacity as "lower" and commercial as "upper". The order can reverse. A numerical interval can be shown solely as the range between two *different methods*, with both method labels retained.

10. Full defaults and mandatory classifications

10.1 Baseline fictional company

Variable	v1 default	Classification
Industry	B2B Software / SaaS	Illustrative

Variable	v1 default	Classification
Previous-year revenue R_0	£50,000,000	Fictional company
BAU revenue growth b	0%	Illustrative
Incremental contribution margin m	75%	Illustrative gross-margin-derived **proxy**
Engineers N	100 FTE	Fictional company
Workforce location	UK 100%	Fictional company
Seniority mix	Junior 20%; mid 50%; senior 30%	Fictional company
Loaded costs	UK £52k / £78k / £117k	Illustrative, not benchmarked
Annual licensed share	70%	Illustrative
Annual effective adoption	60%	Illustrative
Monthly licence price	£25 per seat	Illustrative generic tool
Additional usage/tokens	£15 / active user / month	Illustrative generic tool
Implementation costs I	£0	Excluded by default; disclose
Eligible work share f	50%	Illustrative
Net eligible-task time reduction s	20%	Illustrative
Engineering redeployment u	50%	Illustrative
Engineering commercial response e	65%	Illustrative and **not causal estimate**
Positive commercial diffusion loss d	25%	Illustrative
Downside commercial transmission t_-	100%	Conservative policy assumption
Appraisal period	Single 12-month year	Fixed v1 scope
Reporting currency	GBP	User-changeable

Default values are preloaded for an instant demonstration. **A user must be asked to confirm or replace company-specific inputs before claiming their own company is represented.** All model outputs include the labels *illustrative estimate* and *not an observed financial result* until evidence is provided.

10.2 Industry presets — unvalidated demonstration values

Industry	Incremental margin proxy	Engineering response e	Positive diffusion loss d	Eligible work f
B2B Software / SaaS	75%	65%	25%	50%
IT Consulting / Professional Services	35%	40%	45%	50%
Financial Services	40%	30%	55%	40%
Manufacturing / Industrial	30%	20%	65%	35%
Retail / E-commerce	25%	15%	65%	40%
Other / Custom	40%	30%	50%	40%

The financial-services margin is a **contribution-margin proxy**, not a conventional bank gross margin. Industry selection never silently overrides a manually modified revenue, headcount, location or salary input. An untouched preset-derived assumption can change to the new industry default; manually overridden assumptions stay unchanged unless the user explicitly resets them. Industry submission suggestions are proposals for later curation, not automatic authoritative benchmarks.

10.3 Data provenance requirements

For every preset or reference number store: identifier, value, unit, industry, location/seniority if applicable, classification (fictional / illustrative / evidence-backed / user-supplied), original source URI, source date, review date, reviewer, effective version and notes. An unvalidated illustrative number must have its `validated` status set to false; it must not inherit credibility from cited research about different phenomena.

11. Numerical reference case — independently reproducible

Apply all **default fictional SaaS values** from section 10.

Step	Calculation	Result
Weighted loaded cost	$0.2(£52k) + 0.5(£78k) + 0.3(£117k)$	**£84,500/FTE**

Step	Calculation	Result
Total payroll W	100 × £84,500	**£8,450,000**
Licences	70 × £25 × 12	**£21,000**
Additional usage	60 × £15 × 12	**£10,800**
Total AI cost C	£21,000 + £10,800	**£31,800**
Released time r	50% × 20%	**10.0000%**
Theoretical active-user output uplift g	10% / (1−10%)	**11.1111%**
Average realised engineering output uplift	60% × 50% × 11.1111%	**3.3333%**
Cost-equivalent capacity benefit B_C	£8.45m × 60% × 50% × 10%	**£253,500**
Capacity net value V_C	£253,500 − £31,800	**£221,700**
Capacity economic ROI	£221,700 / £31,800	**697.17%**
Incremental revenue scenario ΔR	£50m × 65% × 3.3333% × 75% retention	**£812,500**
Commercial contribution B_R	£812,500 × 75% incremental margin	**£609,375**
Commercial net value V_R	£609,375 − £31,800	**£577,575**
Commercial scenario ROI	£577,575 / £31,800	**1,816.27%**

This is the **normative regression fixture**. Note that the earlier specification's £1,218,750 commercial benefit **did not apply the shared engineering redeployment factor**; it is superseded and must not be used as a reference output.

11.1 Single-variable productivity sensitivity

Other inputs held at default values. Scenarios vary **eligible-task time reduction**, not directly the overall output uplift.

Scenario	Eligible-task time reduction	Released overall time	Realised output uplift
Conservative	10%	5%	1.5789%
Base	20%	10%	3.3333%
Optimistic	30%	15%	5.2941%

Scenario	Capacity benefit	Commercial contribution benefit	Capacity / commercial ROI
Conservative	£126,750	£288,651.32	298.58% / 807.71%

Scenario	Capacity benefit	Commercial contribution benefit	Capacity / commercial ROI
Base	£253,500	£609,375	697.17% / 1,816.27%
Optimistic	£380,250	£967,830.88	1,095.75% / 2,943.49%

The three cases are **sensitivity tests**, not statistically derived confidence bounds or probability distributions. Do not label them 5th/50th/95th percentiles.

11.2 Valuation reversal threshold

For uniform positive productivity/adoption/utilisation, commercial benefit exceeds capacity benefit when:

$$R > \frac{W\,r}{m\,e\,g\,(1-d)}.$$

Under all other default parameters, that crossover is **£20.8 million** annual BAU revenue. Below it, commercial value is numerically lower than capacity value. This demonstrates why neither method has a fixed upper/lower status.

11.3 Negative stress case

Keep other defaults and set eligible-task **time reduction to −20%**, giving $r = -10\%$ and $g = -9.0909\%$. Apply full downside transmission, and do not use positive redeployment or positive diffusion discounts:

- Capacity downside benefit: **−£507,000**.
- Commercial scenario contribution: approximately **−£1,329,545.45**.
- Tool cost remains **£31,800**.

The negative outcomes are stress-test illustrations, not forecasts; discontinuing a failing rollout may mitigate them.

12. Calculation state, UI, saving and exports

12.1 User flow

Anonymous users can configure a calculation and see updated results immediately. A simple mode shows industry, previous-year revenue, margin, engineering FTE, workforce mix, tool cost and headline productivity. Advanced mode exposes the cohort matrix, input-mode choice, adoption, engineering redeployment, commercial responsiveness, diffusion, margin definitions, exclusions and scenario sensitivity. Saving or retrieving calculations requires a simple account. Public sharing is opt-in.

12.2 Result hierarchy

Display all of the following, with units and qualifying labels:

1. Incremental annual AI expenditure, broken down into licence/usage/implementation.
2. AI-eligible work and measured net productivity assumption.
3. Time saved by active engineers, total workforce released time, and **realised comparable output uplift**.
4. Capacity-cost-equivalent benefit, net economic value and economic ROI.
5. Scenario incremental revenue, incremental contribution benefit, net value and scenario ROI.
6. Sensitivity results, assumption provenance, exclusions and model version.

No screen may say that ROI is "guaranteed", "cash saved" or "forecast revenue" without the corresponding evidence or adjusted cash-flow model. High sensitivity should be presented as a risk, not concealed behind a single large ROI metric.

12.3 Versioned immutable snapshots

Save both the original inputs and all derived results, financial-model implementation identifier, methodology version, industry preset version, compensation table version, exchange-rate snapshot, assumption classifications, overrides, analysis date and output schema version. A saved report reopens its **original immutable result**. Recalculation under a new model creates a new explicitly identified calculation version and diff, never silently overwrites historical results.

12.4 Sharing and confidentiality

Private by default. Shared reports are redacted, immutable snapshots identified by unguessable share tokens, with user-controlled field visibility, expiry and revocation. Shared previews and generated social cards reveal only explicitly selected facts. A visitor may duplicate *permitted public assumptions* without obtaining the owner's private inputs or identity. Exported PDF/CSV results include the methodology version, assumptions and model limitations.

13. Validation, invariants and acceptance tests

13.1 Input validation

- All engineering FTEs nonnegative and $N = \sum_i n_i$ exactly within specified numeric precision.
- Location and seniority proportions total 100% when $N > 0$; users may enter fractional average FTEs.
- All cohort currencies have an explicit rate to reporting currency; refuse silent FX guessing.

- Effective active engineers cannot exceed licensed seats under a paid-seat licence model.
- $f, a, u, d, e, t_ \in [0,1]$.
- $s \in [-0.5,0.8]$, $r = fs < 1$, $g = r/(1-r) > -1$; for direct g, validate implied r.
- $R_0 \geq 0$, $1 + b \geq 0$; allow explicit negative m but warn on interpretation.
- $C \geq 0$; if $C = 0$, no numerical ROI.
- Invalid or missing inputs show an error; never quietly coerce NaN or infinity to zero.

13.2 Required model invariant tests

Test ID	Change / condition	Required observable behaviour
T01	All default inputs	All outputs in §11 match, at display rounding
T02	$a = 0$	Benefits 0; paid licences can still create cost; ROI −100% if $C > 0$
T03	$u = 0$ with positive productivity	Both values 0; no commercial uplift
T04	$f = 0$ or $s = 0$	Both economic benefits 0
T05	$d = 100\%$ and positive productivity	Commercial positive benefit 0; capacity unaffected
T06	$e = 0$	Commercial benefit 0; capacity unaffected
T07	Increase licensed seats only	Cost rises, gross benefits unchanged unless adoption changes
T08	Increase token price only	Cost rises, gross benefits unchanged
T09	Change workforce salary only	Capacity valuation changes; commercial benefit stays the same if N and other inputs fixed
T10	Change contribution margin only	Commercial benefit changes; capacity value unchanged
T11	Change industry after custom override	Custom values preserved; untouched preset values update
T12	$C = 0$	ROI reported ‘not defined’; no divide-by-zero

Test ID	Change / condition	Required observable behaviour
T13	Negative task productivity	Use adverse channel; no positive u, d discounts applied
T14	Change FX rate	Converted payroll and capacity value change consistently; original salary preserved
T15	$N = 0$	Benefits 0; prohibit positive active seats; correctly handle any independent implementation charge
T16	Input-mode equivalence	A task-time case and exactly equivalent direct-output case give equal results
T17	Reverse range case	Display independent method labels and no preassigned upper/lower bounds
T18	Restore old snapshot	Historical outputs unchanged after model/preset updates
T19	Share link revoked	No further unauthorised access to the report
T20	Negative m	Signs and user warnings behave coherently; no hidden positivity clamp

13.3 Property-based tests

Within the positive, otherwise fixed parameter region: greater a, u, f or s (where nonnegative) should not reduce positive benefits; raising d must not increase positive commercial benefit; increasing AI cost must not increase either ROI with positive C and fixed benefits. Holding N fixed, compensation changes must not independently change scenario revenue benefit. Exact algebraic input-mode equivalence should be tested. Monetary calculations must not round until presentation; use an appropriate decimal arithmetic approach for currency values and deterministic reference fixtures.

13.4 Auditability

For each output, generate a **calculation trace** with the equation identifier, input variable values, intermediate results, units, rounding policy, data provenance, decision on defaults and any validation warnings. Provide structured machine-readable output as well as

PDF/CSV. Independent reviewers should be able to reproduce every result from the snapshot without querying an external pricing or FX API.

14. Known limitations and model risk register

ID	Limitation / risk	Materiality	Containment
L01	Historic revenue does not causally identify marginal sales unlocked by engineering	**Critical**	e explicitly a scenario response assumption; sensitivity and warnings
L02	Industry-specific e, d, m and f values are not empirically benchmarked	**Critical**	Label as illustrative; external calibration before 'benchmark' claims
L03	The two valuations use the **same underlying output gain**	**Critical**	Never add values; parallel alternatives only
L04	Time saved is not necessarily avoided payroll or new cash flow	High	Label capacity as economic opportunity cost, not savings
L05	u and d can overlap if users give them vague meanings	High	Require strict upstream redeployment versus downstream commercial definitions
L06	Product output is not homogenous; feature value and customer demand vary	High	Scenario sensitivity; model is an approximation
L07	Average-to-marginal scaling assumes locally linear revenue response	**Critical**	Explain linear proxy; warn at high uplift/material revenue change
L08	Single-year analysis ignores lagged revenue, cohorts	High	Clearly restrict to annual scenario; multi-year model deferred

ID	Limitation / risk	Materiality	Containment
	and recurring-revenue persistence		
L09	Revenue/employee and compensation baselines may be different periods	Medium	Require annual average FTE and explicit periods
L10	Tool licences exclude hidden implementation, platform or governance costs if users choose $I = 0$	High	Disclose omitted cost; user-adjustable I
L11	Net productivity may include unmeasured rework, quality degradation or concurrent AI agents	High	Require quality-adjusted end-to-end productivity assumptions
L12	Negative outcome convention is deliberately conservative and not symmetric	Medium	Separate loss transmission and show stress-case disclosure
L13	Adoption may be correlated with productivity, location or seniority	Medium	Cohort model; do not assume independence if data exists
L14	Large linear uplift may exceed market absorption or organisational capacity	High	Warn at >25% modelled incremental revenue / baseline revenue, allow override
L15	Inflation, taxes, accounting treatment and cash timing are not modelled	Medium	Annual economic-value scope; no NPV/payback in v1
L16	Comparison against existing AI tooling invalid without	**Critical**	Restrict to new tool versus no-tool BAU

ID	Limitation / risk	Materiality	Containment
	incremental baseline adjustment		
L17	*e* and *d* are not separately identifiable from ROI outcomes	High	Show combined commercial response and source each assumption independently
L18	Salary reference values, FX and licence figures can become outdated	Medium	Version, date, cite and mark unvalidated values
L19	ROI may be numerically large due to low subscription costs, not high-confidence revenue attribution	High	Show monetary values, assumptions and sensitivities as prominently as ROI
L20	Positive- and negative-productivity values cannot be treated as a confidence interval	Medium	Explicit scenario labels; no fake probability or precision

Residual risk: Even after fixing internal inconsistencies, an independently reproduced result is **not a validated causal estimate**. A finance-led use case requires firm-specific measured developer performance, actual vendor spend, validated incremental contribution margins, and preferably observed revenue or customer-outcome evidence.

15. Evidential standards and calibration plan

15.1 Priority evidence for a real organisation

1. **Incremental tool cost:** actual contract, seat count, overages, training/integration and baseline spend.
2. **Net productivity:** comparable before/after task data with controls where feasible, accounting for review, rework, testing and output quality; or a carefully designed experiment.
3. **Adoption and utilisation:** telemetry and observed capacity redeployment, not licences issued or self-reported gains alone.
4. **Commercial uplift:** product-release / sales pipeline / customer-demand mechanisms linking additional output to attributable incremental revenue.

5. **Contribution margin:** actual incremental variable cost analysis rather than indiscriminate use of company gross margin.
6. **Counterfactual:** realistic BAU progress, existing tools and expected normal growth separated from attributed AI outcomes.

15.2 Research context (not validation of defaults)

The 2025 DORA research argues that AI can amplify strengths and bottlenecks in software-delivery organisations; it does not establish a single transfer rate from faster engineering to additional sales. METR's 2025 experiment observed an AI-associated slowdown in its experienced-open-source-developer setting, while a 2026 METR update describes methodological selection limitations in its newer measurements. These findings support **measuring net productivity and allowing negative outcomes**, not importing the measured slowdown or any universal uplift as the default. HM Treasury's 2026 Green Book discusses appraisal against a credible BAU and the distinction between outcomes that would have happened anyway and additional outcomes; the calculator borrows that appraisal discipline but does **not claim Green Book compliance**.

15.3 Calibrating the industry library

Before representing a preset as evidence-backed, obtain sources of consistent scope for each parameter, document estimation method, geography and recency, separate descriptive statistics from causally estimated effects, assign review owners and record uncertainty. **Engineering commercial response and diffusion will usually require user-specific judgement** rather than a single defensible industry figure. Until this work is complete, all industry seeds remain labelled *illustrative scenario assumptions*.

16. Scope decisions and change log

16.1 In scope — financial engine v1

- One-year, no-AI BAU versus incremental AI tool; positive and negative net productivity.
- Licence, overage and optional deployment costs.
- Seniority/location fully loaded workforce matrix and dated FX conversion.
- Task time-saving mode and equivalent end-to-end output-uplift mode.
- Distinct adoption, productive redeployment and downstream commercial diffusion.
- Two mutually exclusive monetary methods, separate net benefits and ROIs.
- Industry/user defaults, full overrides, scenario/sensitivity analysis, provenance, testing and immutable audit records.

16.2 Deferred until explicitly re-specified

- Multi-year adoption curves, time lag, retention, revenue persistence and NPV/payback.
- Migration from incumbent AI tools with negative incremental licence cost.

- Verified cash payroll savings and hard-cost takeout; addition of genuinely disjoint benefit channels.
- Detailed by-vendor token metering and billing commitments.
- Causal econometric calibration of marginal commercial response.
- Team collaboration, paid plans and enterprise benchmarking.

16.3 Changes from v1.0 to v1

Previous concern	Adopted v1 treatment
Capacity realisation absent from revenue path	Shared **productive redeployment** u applied to both positive benefit channels
Engineering contribution confused with factual profit attribution	Redefined e as a **commercial response proxy**; explicit linear approximation
Gross and incremental contribution margin conflated	Commercial benefit calculated using one **incremental contribution margin** m
Historical contribution pool treated as guaranteed upper bound	Two independent valuations without fixed numerical ordering
Negative productivity discounted by positive-benefit factors	Separate adverse channel and explicit downside transmission t_-
Multi-year mechanics underspecified	One-year analytical scope; future forecasts deferred
Counterfactual unclear	Explicit new deployment vs no-tool baseline and incremental BAU rule
Demonstration defaults could appear authoritative	Mandatory **unvalidated** classification and provenance
Saved results could change under code updates	Full snapshot + versioned methodology and calculation trace

17. Independent audit protocol and acceptance gate

An independent reviewer should answer the following before sign-off:

Review area	Required reviewer evidence / decision
Mathematical integrity	Re-derive each equation; recompute default, sensitivity, zero and negative fixtures
Dimensional consistency	Confirm counts, percentages, time and monetary units are coherent
Counterfactual integrity	Verify that adoption, costs and productivity are incremental versus BAU

Review area	Required reviewer evidence / decision
Economic classification	Confirm capacity is not presented as cash and commercial value is not guaranteed profit
Double counting	Confirm productivity, u, d, margin and optional cost categories do not overlap
Parameter meaning	Confirm e and d are comprehensible, distinguishable and properly qualified
Financial reporting	Review margin definition, baseline periods, loaded cost conventions and currency conversion
Assumption governance	Review source flags, industry presets, salary values and override rules
Numerical implementation	Test reference examples, rounding, boundaries, negative returns and division by zero
Results disclosure	Check PDF, CSV, shared reports, warnings and privacy controls
Model risk	Assess whether sensitivity adequately communicates risk of extreme commercial output
Release decision	Record audit verdict, unresolved issues, remediation owner and sign-off date

Proposed release gate: engineering may implement v1 as a transparent **illustrative scenario calculator** after mathematical review. Do not market its industry presets as independently benchmarked nor its revenue-linked output as a validated forecast until the model's causal assumptions and reference values have been externally supported. This document is a specification for audit, **not itself an audit opinion**.

18. Primary references and interpretation

1. **DORA / Google (2025)**, *State of AI-assisted Software Development*. Research on organisational context, delivery bottlenecks and AI effects. https://dora.dev/research/2025/dora-report/.
2. **METR (2025)**, Becker et al., *Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity*. A study of particular developers, projects and tool generations; not a universal benchmark. https://arxiv.org/abs/2507.09089.
3. **METR (24 February 2026)**, *We are Changing our Developer Productivity Experiment Design*. Explanation of selection effects and uncertainty in newer measurement. https://metr.org/blog/2026-02-24-uplift-update/.
4. **HM Treasury (28 October 2026)**, *The Green Book (2026)*. General public-sector appraisal guidance, used solely to illustrate disciplined BAU/additionality thinking;

this private SaaS tool is not certified against it. https://www.gov.uk/government/publications/the-green-book--2/the-green-book-2026.

End of controlled draft — Version 1