



What 500 Megaprojects Show: The Evidence Base Behind the Lab's Figures

Coverage, definitions and headline measurements from the Megaproject Lab case base

ABSTRACT

Megaproject Lab is an independent research initiative focused on understanding how major infrastructure projects evolve through the interaction of political and technical dynamics over time. The initiative starts from the premise that megaproject risks are not static events, but evolving conditions shaped by decisions, incentives, uncertainty, timing, stakeholder behavior and institutional pressures across the project lifecycle.

This note documents the evidence base behind the figures the Lab publishes. It describes the case base and its coverage, defines how outcomes are measured, and reports the headline measurements with their sample sizes: cost and schedule performance, economic performance against appraisal, and one worked example on the timing of public announcements relative to feasibility work.

The note is descriptive. It reports frequencies and central values together with the definitions that produce them, and it states the caveats a reader should carry. The calibrated model built on this base, with its rules and parameters, is documented separately and is outside the scope of this note.

1. The Case Base

The case base contains 500 major infrastructure projects, documented case by case from public evidence. Execution of the projects included began between 1987 and 2023, and in all but one case after 2000. The portfolio spans about one hundred countries and the main infrastructure sectors, including transport, energy, water and social infrastructure.

The unit of analysis is a bounded project or programme. Each case states its working boundary before coding begins: whether it covers a full programme or a defined phase, which associated works are included because they materially affect cost, schedule, readiness or benefits, and which later expansions are excluded. Multinational projects are coded once, with every participating country identified. The boundary decision is recorded with the case, so that every score and outcome refers to a stated scope.

Every case follows the same coding method. A written codebook defines 65 variables covering technical, economic, social, environmental, institutional and political conditions, each scored from 0 to 10 against predefined anchors that describe how strongly the condition appears in the evidence. Each case also records the risks that materialized, the causal mechanisms observed, the quantified outcomes, and the sources that support every entry. A structured human review gate precedes insertion into the base.

The resulting volume, as of the version used for this note: 22,471 variable scores, 36,184 stage-specific variable observations, 4,471 observed risks, 3,219 observed mechanisms, and 2,518 outcome observations with a quantified deviation. The evidence layer holds 23,433 curated sources, with a median of fifty sources per case.

All figures in this note were computed on version 1.0 of the base in September 2026. The computation scripts and per-figure definitions are kept on file.

2. How the Base Is Built

Each case starts from a scouting pass over the public record: government decisions, approval documents, budget and parliamentary records, multilateral project documents, audit reports, regulatory filings and established press. Sources are curated before coding, with a stated preference for official, multilateral and audit material, and each source is kept on record together with the claims it supports.

Coding then proceeds in stages against the written codebook: the chronology of lifecycle events first, then variable scores with their evidence, then observed risks and mechanisms, then quantified outcomes. Every entry cites the sources behind it. Where the evidence is thin, the entry records lower confidence in place of a guessed score.

A structured human review closes the process. Each case reaches the base through a review gate where the coding is examined and approved before insertion, and every insertion is versioned, with a dated backup of the previous

state of the base. The figures in any Lab publication therefore point to a fixed and reproducible version of the base.

3. How Outcomes Are Measured

An outcome observation is a deviation documented in the public record between a realized value and the approved baseline it is compared against in that record: realized cost against approved cost, realized completion against approved completion. Each observation is stored as a percentage deviation.

When a case carries several observations of the same type, the case takes the median of its observations, so that heavily reported projects do not weigh more than sparsely reported ones. Deviations outside a window of minus 90 percent to plus 3,000 percent are excluded as implausible or non-comparable.

A project counts as over budget, or late, when its per-case deviation exceeds plus 5 percent. The threshold separates real deviations from rounding and re-statement noise. Central values are reported as medians, and the corresponding means are disclosed alongside them. The means run substantially higher because a minority of extreme cases pulls them up, which is itself part of the record.

Baselines differ across jurisdictions in price treatment and re-approval practice. The figures aggregate deviations as documented, and that heterogeneity is a limitation carried explicitly.

4. Cost and Schedule

Of the projects in the base, 343 carry usable cost outcomes. Sixty-four percent of them ran over their approved cost. Among those that ran over, the median overrun is 47 percent, and the mean is 81 percent. Thirty percent of all projects with cost outcomes exceeded their approved cost by half or more.

Schedule outcomes are usable for 287 projects. Fifty-four percent finished late. Among the late projects, the median delay is 61 percent of the approved schedule, and the mean is 105 percent. Thirty-one percent of all projects with schedule outcomes ran half again their approved schedule or longer.

The far end of the distribution is thick. Sixteen percent of projects with cost outcomes ended at or above twice their approved cost, and 17 percent of projects with schedule outcomes took at least twice their approved time.

5. Economic Performance Against Appraisal

Economic outcomes are the scarcest in public records. Fifty-one projects in the base carry a measurable comparison between realized economic performance and the appraisal expectation. Of those, eight in ten performed below the appraisal figure by more than 5 percent.

The sample is small, and it is reported with its size for that reason. Projects that publish an ex post economic evaluation may differ from those that do not, and the direction of that selection is unknown.

6. Demand Against Forecast

Demand outcomes compare realized demand with the forecast used at appraisal. Seventy projects in the base carry a usable comparison, and 54 percent of them realized demand below forecast by more than 5 percent.

The note reports the frequency alone for this outcome. The underlying observations mix early ramp-up years with mature operation, and a central value would average across that mix. Refining that comparison is part of the ongoing work on the base.

7. A Worked Example: Announcement Timing

One variable in the codebook records whether a project was announced publicly before its feasibility work was complete. The variable is scored from 0 to 10 against written anchors, from no evidence of premature announcement to a formal public commitment made well ahead of the studies.

Two groups make the comparison. Projects where the condition is strongly present, scoring 7 or above, and projects where it is weakly present or absent, scoring 4 or below. Among projects with cost outcomes, the strongly present group counts 111 cases with a median overrun of 37 percent, while the weakly present group counts 12 cases with a median deviation of minus 1 percent, meaning no median overrun. Among projects with schedule outcomes, the groups count 87 and 11 cases, with median delays of 49 and 5 percent.

Three caveats apply. The weakly present group is small, and the contrast is reported with its sample sizes. The association is descriptive, and no causal claim follows from it. The two groups may differ in composition by sector, region or scale, and the comparison does not adjust for that.

The figure is published because of what it illustrates. A condition that appears in no standard risk matrix separates

the two groups more sharply than most conditions that do.

8. How Conditions Move Between Stages

The base observes variables by lifecycle stage, and half of the variable-and-case pairs carry observations at two stages or more, 11,396 trajectories in total. Among them, 72 percent change value between stages, and one in three moves across scoring bands. The documented conditions of a megaproject move over its life.

The order of movement is also visible. Where two risks appear in the same case with dated materializations, stable orderings emerge. Stakeholder conflict is observed before cost overruns in 83 percent of such pairs, and design-maturity problems before contractual stress in 87 percent. Preparation-stage and stakeholder conditions tend to appear first, contractual stress next, and cost, economic and operational consequences later.

9. What the Base Does Not Show

Some questions cannot be answered from public records, and the base does not claim to answer them. The effectiveness of mitigation strategies has no counterfactual: where a project acted and a risk did not materialize, the record cannot show what would have happened otherwise. The base documents what projects did, while the value of what they prevented remains outside measurement.

Internal project registers are also outside the record. Contract-level claims, internal schedule variances, unpublished risk registers and internal appraisals appear in the base only when a public document discloses them. The base measures the documented life of projects, defined as what a diligent outside reader could establish from the public record.

Projects that disappear quietly are the hardest to document. A cancelled study or a shelved proposal leaves fewer traces than a completed asset, so early terminations are underrepresented relative to their real frequency, and the headline figures should be read with that limitation in mind.

10. How to Read These Figures

Two further limitations frame the numbers. Baselines are taken as documented, and jurisdictions differ in how they set and restate them. The figures are descriptive statistics of a curated portfolio, and carry no causal interpretation by themselves.

Within those limits, the figures are reproducible. Each one carries a definition, a sample size and a computation script,

and the version of the base that produced it is fixed and dated. Readers who need a figure not reported here, or a breakdown by sector or region, can write to the Lab.

Suggested citation: Megaproject Lab, "What 500 Megaprojects Show: The Evidence Base Behind the Lab's Figures", Working Note N° 2, version 1.0, September 2026. megaprojectlab.com.

Megaproject Lab · megaprojectlab.com