OXFORD GLOBAL PROJECTS

Reference Class Forecast for the for York Place to Newhaven Project

BACKGROUND AND PURPOSE OF THIS BRIEFING NOTE

Edinburgh City Council has commissioned Oxford Global Project to conduct a Reference Class Forecast to estimate the cost and schedule risk of the planned extension of the Edinburgh tram from York Place to Newhaven. The Reference Class Forecast was prepared by our team under the leadership of Dr. Alexander Budzier and under the senior supervision of Prof. Bent Flyvbjerg.

METHODOLOGY

Reference class forecasting is the method of predicting the future, through looking at similar past situations and their outcomes. Reference class forecasting predicts the outcome of a planned action based on actual outcomes in a reference class of similar actions to that being forecast. The theories behind reference class forecasting were developed by Daniel Kahneman and Amos Tversky. The theoretical work helped Kahneman win the Nobel Prize in Economics.

Kahneman and Tversky (1979) found that human judgment is generally optimistic due to overconfidence and insufficient consideration of distributional information about outcomes. Therefore, people tend to underestimate the costs, completion times, and risks of planned actions, whereas they tend to overestimate the benefits of those same actions. Such error is caused by actors taking an "inside view," where focus is on the constituents of the specific planned action instead of on the actual outcomes of similar ventures that have already been completed.

Kahneman and Tversky (1979) concluded that disregard of distributional information, that is, risk, is perhaps the major source of error in forecasting. On that basis they recommended that forecasters "should therefore make every effort to frame the forecasting problem so as to facilitate utilizing all the distributional information that is available". Using distributional information from previous ventures similar to the one being forecast is called taking an "outside view". Reference class forecasting is a method for taking an outside view on planned actions.

OFFICIAL GUIDANCE ON OPTIMISM BIAS AND REFERENCE CLASS FORECASTING

Reference Class Forecasting is the method to establish the “Optimism Bias Uplift” required by HM Treasury’s Greenbook:

> *“Optimism bias is a form of reference class forecasting which predicts future outcomes based on the outcomes for a group of similar past projects. It is important to note that adjustments for optimism bias are not the same as financial contingency.” (HMT Greenbook, 2018, paragraph 5.46, p. 30).*

The last point is important, because HMT recognizes that while some questions raised during project appraisal might require a conservative estimate of risk (e.g. If an extreme downside scenario would occur, would this bankrupt the delivery vehicle?), however, it is not always the best use of resources to actually allocate for a high level of risk, e.g. an estimate at 5% accepted chance of an overrun (P95).

Further the Greenbook states:

> *“Ideally adjustments should be based on an organisation’s own evidence base for historic levels of optimism bias.” (HMT Greenbook 2018, paragraph 5.45, p. 30).*

For this, we have used the best available data, in the form of 152 light rail transport projects, from our academically peer-reviewed project database.

The process of establishing a reference class forecast is further detailed in the UK Department for Transport’s guidance document titled “Procedures for Dealing with Optimism Bias in Transport Planning” of 2004. The guidance outlines the three-step approach to formulate a Reference Class Forecast to establish the required level of Optimism Bias and risk:

> *“Taking an outside view requires the following steps for the individual project:*
>
> - *Identification of a relevant reference class of past projects*
>   *The key is here that the class is broad enough to be statistically meaningful but narrow enough to be truly comparable with the specific project.*
>
> - *Establishing a probability distribution for the selected reference class*
>   *This requires access to credible data on cost increases (or time schedule delays or benefit shortfalls if these are the key parameter) on a sufficient number of projects within the*

reference class to make statistically meaningful conclusions (normally at least 10).

- *Placing the specific project at an appropriate point in the reference class distribution*
  This step has an element of intuitive assessment and is therefore liable to optimism bias." (DfT, 2004, p. 9)

The remainder of this briefing note applies this three-step process to the Edinburgh Tram extension project.

STEP 1 – IDENTIFICATION OF THE RELEVANT REFERENCE CLASS OF PAST PROJECTS

The reference class was established through statistical analysis of past, completed light rail projects.

For this analysis the variable in question are cost and schedule overrun.

Cost overrun is measured as *Actual Cost / Estimated Cost – 1*, where estimated cost is measured at the full business case stage and actual cost at project completion. Cost overrun is measured in real terms, i.e. inflation has been removed.

Schedule overrun is measured as Actual Schedule / Estimated Schedule – 1, where estimated schedule is measured from the approval of the full business case, i.e. the date of decision to build, to the planned date of completion. Actual schedule is measured as the time from full business case approval to scheme completion.

The analysis tested whether the 152 international light rail projects for which data were available are comparable to the planned extension project.

The analysis concluded that only extension projects should be included because they have a statistically significantly lower cost and schedule risk as new build light rail projects. A full list of analyses carried out and their findings are included in Annex A.

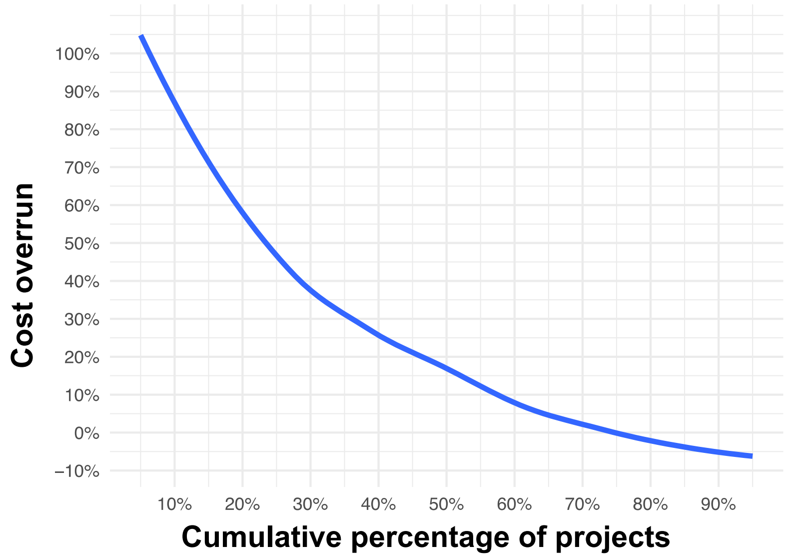
This resulted in a reference class of 89 past, completed light rail extension projects.

STEP 2 – ESTABLISHING A PROBABILITY DISTRIBUTION FOR THE SELECTED REFERENCE CLASS

The cumulative distribution was calculated for these 89 projects. For this, the projects are ordered by size of overrun and then their cumulative share in the sample is calculated.

Figure 1 shows the probability distribution of cost overrun. The data show that half the projects had a cost overrun of more than approximately 17% and half of the projects had a cost overrun of less than approximately 17%.

Figure 1 Probability distribution (here cumulative) of cost overrun for the selected reference class



The data are then used to establish the required optimism bias uplift as a function of the acceptable chance of cost overrun, i.e. the risk appetite of decision makers.

Figure 2 shows that based on this reference class forecast a decision maker willing to accept a 50% chance of a cost overrun (P50) should uplift their project cost estimate by 17%.

A more conservative decision maker might only be willing to accept a 20% chance of cost overrun that would require a 57% uplift on the estimated cost.

Figure 2 Establishing the required optimism bias uplift as a function of the acceptable chance of cost overrun

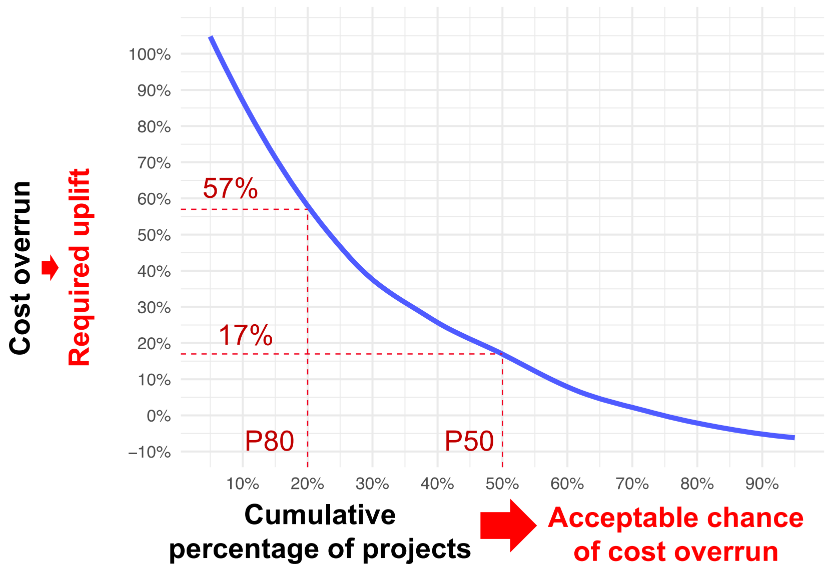


Table 1 shows the results for the reference class forecasts for cost and schedule overruns of past, completed light rail transport projects.

Table 1 Results of the reference class

	Expected overrun (Mean)	**50% acceptable chance of overrun (P50)**	**20% acceptable chance of overrun (P80)**	**10% acceptable chance of overrun (P90)**	**5% acceptable chance of overrun (P95)**
Cost risk estimate (%)	32%	17%	57%	86%	104%
Cost risk estimate (GBP mio)	52.5	27.9	93.5	141.1	170.7
Project cost estimate inclusive of risk and optimism bias (GBP mio)	216.6	192.0	257.6	305.2	334.8
Schedule risk estimate (%)	25%	3%	35%	65%	82%
Schedule risk estimate (delays in months from FBC approval to completion)	13	2	18	33	42
Project completion date	2024-06-07	2023-07-04	2024-11-08	2026-02-13	2026-11-01

STEP 3 – PLACING THE SPECIFIC PROJECT AT AN APPROPRIATE POINT IN THE REFERENCE CLASS DISTRIBUTION

The final step places the extension project in the reference class distribution.

Currently the extension project has made risk provisions in the form of contingency of GBP 43.2 million[1]. This is equivalent to an uplift of approximately 26% on top of the base cost estimate[2]. This 26% provision for risk is equivalent to a 39% acceptable chance of a cost overrun (or a P61), i.e. this level of risk provision has only been exceeded by 39% of past projects and 61% of past projects stayed within this envelope of funding.

1 This includes 31.9 million from the quantitative risk analysis of the identified risk in the risk register of the project and 11.3 million for unidentified risk, i.e. through the generic 6% uplift.

2 This includes 5.5 million of development cost to date, 156.7 millions of capital cost to completion and 1.9 million as support for business.

This acceptable chance of a cost overrun is similar to the funding of other recent transport projects in the UK, for example High Speed 2, which is funded at 33% acceptable chance of a cost overrun.

However, the methodology assumes that the project under consideration is no more or less risky than the past, completed projects in the reference class.

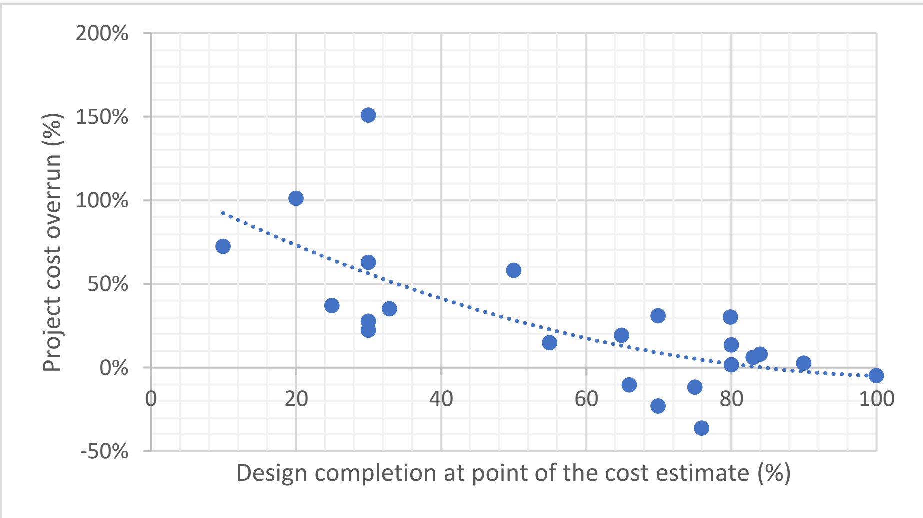
The official guidance suggests that empirical evidence might exist that provides a strong argument that allows to position a project as more or less risky than the reference class.

For this we have analysed the causes of cost overrun in the reference class. The key causes were:

- Low design maturity at the decision to build;
- Scope increases (e.g. adding additional stations, extending the planned alignment, adding additional depots);
- Unforeseen ground conditions: extend of utilities to be diverted and weaker than expected soil; and
- Early delays during procurement.

The first cause of overrun is the level of design maturity. Figure 3 shows the common link between the design maturity at the time when project cost were estimated at the final decision to build and the resulting cost overrun. This study, based on Canadian projects, concluded that increasing design maturity reduces the size of project cost overrun.

Figure 3 Relationship between design maturity and cost overrun (Jargeas 2015)



The second cause is linked to the decision-making process, where cost overrun is caused by scope additions. In light rail transport schemes these scope additions are increasing the number of stations, adding depots and increasing the number of rolling stock.

The third cause are unforeseen ground conditions. It is highly common that the extent of required utility diversions is underestimated. Another commonly unforeseen ground condition is the strength of the soil, which is often found too weak to support the weight of trams.

The fourth cause of cost overrun are early delays. Typically, these occur already during procurement, where the time to conclude the procurement phase is underestimated which results in delays and cost overrun.

Table 2 lists the most common causes of cost overrun and the steps the extension project has taken to address these risks.

Table 2 Common causes of cost overrun in light rail transport schemes and steps Edinburgh tram extension project has taken to address these

Common cause of cost overrun in light rail transport schemes	**Steps taken to address this risk**
Low design maturity at full business case stage	Advanced stages of design; reference design already shared during bidding process
Scope increases	Alignment frozen, political commitment achieved to the current alignment
Unforeseen ground conditions: utilities	Part of utilities already diverted as part of the first tram project
Unforeseen ground conditions: soil strength	Geotechnical investigations
Procurement delays	Early contractor involvement; start of negotiations before FBC approval

The project has already taken steps to de-risk the project. While the impact cannot be quantified this provides demonstrable evidence that the project might be less risky than the projects in the reference class.

EVALUATION

The reference class positions the cost overrun risk of the Newhaven Extension project clearly at approximately P60. The current budget envelope is more likely to be sufficient than not.

In addition, when decision makers and stakeholders evaluate whether the project’s risk of cost overrun is within acceptable levels of risk; decision makers and stakeholders should consider the steps taken by the project to identify and address key risks found in the reference class.

This might result in a greater confidence by decision makers and stakeholders that the project will deliver within the budget envelope than the project's position in the reference class suggests; i.e. that the steps taken have made the project less risky than the past, completed projects.

As mentioned above, other projects, e.g. HS2, have recently been approved at a similar level of risk once they were combined with a programme of project enhancements to provide decision makers and stakeholders with increased confidence in the project's ability to deliver above and beyond the project's position in the reference class of historical projects.

In addition, decision makers ought to consider that any additional funding for risk provisions comes at decreasing marginal benefits; i.e. lower risk appetite comes at ever higher cost. This is because the current position at approximately P60 is close to the tail of the risk distribution.

Thus, managing down the tail risk is often seen to be a better use of resources than increasing contingencies. Additional steps to reduce the tail risk, practices known as Black Swan management, could include identification and reduction of project complexities; setting up an early warning system; and further enhancing the project's delivery capabilities and oversight. Actively addressing the tail risk would further increase the likelihood that the project will outperform previous completed projects.

ANNEX A – FULL LIST OF STATISTICAL TESTS TO IDENTIFY THE RELEVANT REFERENCE CLASS

Table 3 Statistical tests carried out to identify the relevant reference class

Selection question	Result	Conclusion
Should projects from different regions be excluded?	No statistically significant cost overrun ($p \geq 0.74$) No statistically significant schedule overrun ($p \geq 0.15$)	LRT projects have similar cost and schedule overruns across the world. All projects should be considered in the reference class.
Should projects with different physical size (i.e. length in km) be excluded?	No statistically significant cost overrun ($p = 0.2846$) No statistically significant schedule overrun ($p = 0.5076$)	LRT projects of different length (in km, as built) have similar cost and schedule overruns. All projects should be considered in the reference class.
Should projects with an older date of decision to build be excluded?	No statistically significant cost overrun ($p = 0.8346$) No statistically significant schedule overrun ($p = 0.1317$)	LRT projects with different dates of the decision to build have similar cost and schedule overruns. All projects should be considered in the reference class.
Should different project types be excluded?	Cost overruns of new build projects are statistically significantly different from extensions ($p = 0.05$) Schedule overruns of new build projects are statistically significantly different from extensions ($p < 0.001$)	New build projects are statistically significantly different from extensions. Upgrade projects, however, are not. Extension and upgrade LRT projects should be considered in the reference class.