german cooperation
DEUTSCHE ZUSAMMENARBEIT

Implemented by:
giz Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

TCU
FEDERAL COURT OF ACCOUNTS

OLACEFS
ORGANIZACIÓN LATINOAMERICANA Y DEL CARIBE DE ENTIDADES FISCALIZADORAS SUPERIORES

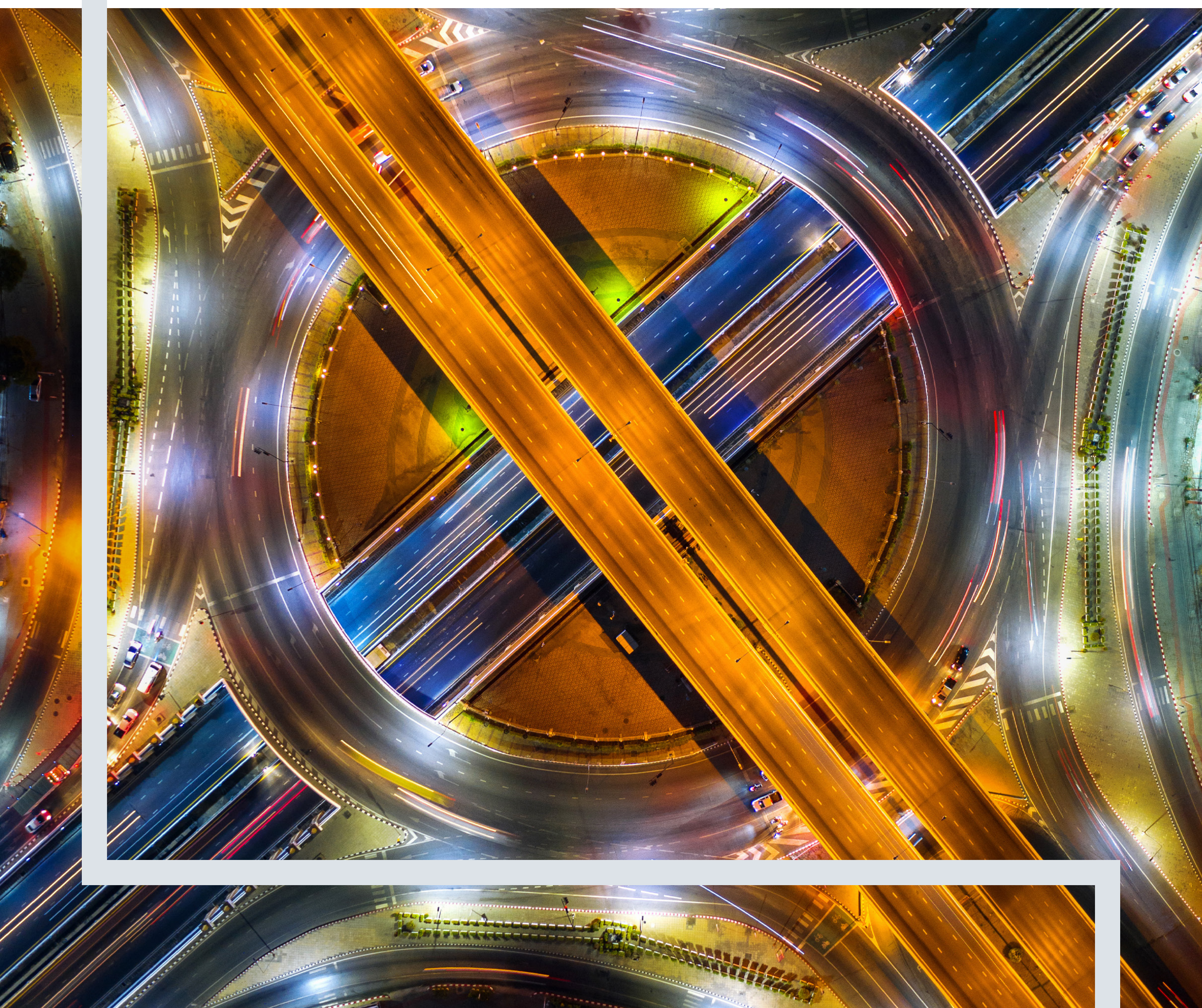
GTInfra
POLÍTICAS Y REGULACIONES DE INFRAESTRUCTURA

REGIONAL PROJECT ANTI CORRUPTION OLACEFS

OXFORD GLOBAL PROJECTS

PRACTICAL REPORT REFERENCE CLASS FORECASTING (RCF)

Methodology for Transport Infrastructure Projects in Brazil



Initiative implemented by the German Cooperation for Sustainable Development, through Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ), in partnership with the Federal Court of Accounts of Brazil (TCU) and the Latin American and Caribbean Organization of Supreme Audit Institutions (OLACEFS), within the scope of the Regional Project Strengthening External Financial Control to Effectively Prevent and Combat Against Corruption.

Prepared by:
Oxford Global Projects UK Limited
John Eccles House Robert Robinson Avenue Oxford Science Park Oxford
United Kingdom
OX4 4GP
info@oxfordglobalprojects.com

Prepared for:
Federal Court of Accounts
SAFS Quadra 4, Lote 1
Brasília - DF - CEP 70042-900
+55 (61) 3527-7222
serint@tcu.gov.br

Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH
SCN quadra 01, Bloco C, Sala 1501
Ed. Brasília, Trade Center
70711-902 Brasília, Brasil
giz-brasilien@giz.de

Practical report on the Reference Class Forecasting (RCF) • Regional Project Strengthening External Financial Control to Effectively Prevent and Combat Against Corruption • October/2023 • Brazil

1. External control 2. Anticorruption 3. Infrastructure 4. Public works 5. Transport Sector 6. Feasibility 7. Optimism bias 8. Strategic Misrepresentation

More information about the Feasibility in Focus Project:

OXFORD GLOBAL PROJECTS

Oxford Global Projects

Andreas Leed

Dirk Pöker

Emma Schubart

Alexander Budzier

October/2023

Table of contents

EXECUTIVE SUMMARY

This report provides high-level benchmarks on the costs and timeframes of Brazilian infrastructure projects relative to similar projects in other countries. It investigates whether there are meaningful and/or statistically significant differences in infrastructure cost and schedule overruns between Brazil and other countries. These quantified outputs of cost and schedule can be used in project appraisal and to forecast the costs, timelines, and risks associated with infrastructure development.

The purpose of this report is to analyze and compare Brazilian performance to the overall distribution of international projects. The findings can be used to motivate and guide further research and investigation into Brazilian infrastructure project performance and its drivers. The report identifies differences in cost performance but does not seek to explain observed differences. Cost overruns, schedule overruns and unit costs of road, rail, bridge and tunnel projects in Brazil have been analyzed. They stem from two different datasets that were obtained by the Federal Court of Accounts (TCU), based on information made available by the Brazilian government bodies of the transport sectors responsible for public works in road and rail. The data was then compared to other international projects stemming from the database of Oxford Global Projects (OGP).

Upon comparison of the Brazilian data with an international benchmark, Brazil seems to have performed significantly worse in terms of schedule overrun. In terms of cost overrun, we found that Brazil performed worse than European, North American and Oceanic countries but better than other South American countries. Note this is not a one-to-one comparison, as the Brazilian data consist of construction contract packages (i.e. partial project cost), while the OGP data is measured at the project level (i.e. full project cost). Due to differences in the datasets as well as limitations found in the Brazilian dataset, the comparison conclusions are not robust.

The report concludes that the datasets collated in Brazil have considerable limitations due to systematic issues of both transparency and detail, since they are grouped by contract and not project. Thus, any analysis of costs and durations based on this data should be interpreted with caution. The report also makes several recommendations on how to address these issues and suggests using international data for benchmarking and reference class forecasting, until the present limitations have been addressed.

1. INTRODUCTION

Public infrastructure investments are important for potential economic growth and productivity. They can provide significant positive spillovers in the economy and can help in the alleviation of poverty and the reduction of income distribution inequality. Therefore, it is generally seen as an important tool in a country's policy toolbox.

But this positive effect can be offset if the investment projects run over budget and/or over schedule. If costs are significantly higher than initially intended, the cost-benefit analysis might shift to a negative net benefit. In fact, research by Flyvbjerg et al. (2003)[1] has shown that investment projects often experience cost overruns and completion delays, and that the positive contribution of infrastructure projects has therefore been questioned.

The Brazilian case is not exceptional in this aspect. Research suggests that around 70% construction projects in Brazil exceeded their budget, with one out of five of them having budget overruns of more than 25% compared to the initial agreement [2]. The causes for such overruns in cost and schedule can be numerous, including inaccurate cost estimates, design modifications, quantity changes, variation orders, political interference, inflation, environmental factors or even malicious behavior such as deception and corruption. Often it is not possible to pin down the exact reason for the non-compliance with the initial agreed costs, which emphasizes the difficulty of performing an accurate cost estimation in the first place.

In general, however, we can identify two main categories for causes of overruns: (1) optimism bias in the planning phase, (2) strategic misrepresentation.

Whereas the first category of causes is common and transcends gender, ethnicity, nationality, expertise, and age[3], the second seems to play a particularly important role in the Brazilian marketplace. According to Transparency International's 2019 Global Corruption Barometer on Latin America and the Caribbean (LAC), the majority of citizens in the region think that corruption in their country has increased over the past 12 months. Only 21% of people in the LAC region have confidence in their government, and 65% think their country's government is run by private interests and serves only selected segments of society.

1 Flyvbjerg B, Bruzelius N and Rothengatter W (2003) Megaprojects and Risk. Cambridge University Press, Cambridge, UK.

2 França, Alda & Haddad, Assed. (2018). Causes of Construction Projects Cost Overrun in Brazil. International Journal of Sustainable Construction Engineering Technology

3 O'Sullivan, Owen P. (2015). The neural basis of always looking on the bright side. Dialogues in Philosophy, Mental and Neuro Sciences

One way to remove optimism bias and strategic misrepresentation in project planning is by adopting top-down forecasting methods that depend on historical data, such as reference class forecasting (RCF). The theories behind RCF were developed by Daniel Kahneman and Amos Tversky. They found that human judgement is generally optimistic owing to overconfidence and insufficient consideration of distributional information about outcomes. When people try to estimate project outcomes, costs, completion times and risks are often underestimated, whereas benefits tend to be overestimated. Such an error is caused by taking an 'inside view', where the attention lies on the specific planned project instead of on the actual historical outcomes of similar previous projects.

RCF adopts a top-down approach which is based on the 'outside view' of past similar projects. It generates cost and time estimates that are based on historical data, account for the systematic underestimation of cost and schedule overrun in projects and correspond to accepted levels of certainty. RCF therefore contributes to improving forecasts by de-biasing estimates, while explicitly considering the risk appetite of decision makers. The method has been widely and successfully implemented by European governments including the United Kingdom and endorsed by several countries.

This document covers the technical consultancy undertaken by Oxford Global Projects (OGP) to analyze the historical costs, durations, cost overruns, and schedule overruns of Brazilian road and railway works (including related walkways, bridges, highways, tunnels, and viaducts) and develop guidance material for using reference class forecasting. This work can further be used as a reference for the development of the method in other regional Supreme Audit Institutions members of OLACEFS.

The study begins by giving an overview of the methods used and the theory behind them. It then continues by describing the data utilized for the analysis. The results of the analysis are subsequently presented, and corresponding conclusions drawn. The report finishes by giving recommendations in order to improve and stimulate future research.

2. PROJECT SCOPE

Supreme Audit Institutions (SAIs) can contribute to improving the transparency of public administration, making risks visible, and creating robust and effective internal controls to contribute to corruption prevention. The Latin American and Caribbean Organization of Supreme Audit Institutions (OLACEFS) is an international, autonomous, independent, non-partisan, and permanent body that has served since 1963 as a forum to promote the exchange of knowledge related to the audit and external control of the government, as well as to foster cooperative and capacity-building relationships among its 22 members.

The German Cooperation for Sustainable Development, through Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, in partnership with the Federal Court of Accounts of Brazil (TCU) and the Latin American and Caribbean Organization of Supreme Audit Institutions (OLACEFS), have been implementing the regional project for Strengthening External Control for the Prevention and Effective Combat of Corruption ("Project") since May 2021. Prepared by the German Federal Ministry for Economic Cooperation and Development (BMZ), the Project aims to achieve more active involvement of Civil Society Organizations (CSOs) in national anti-corruption systems, including the current period of the COVID-19 pandemic.

Historically, the serious problems of technical, economic, and social-environmental feasibility analysis in infrastructure projects in the Brazilian, regional, and global contexts are well known. And this problematic situation remains current, especially concerning economic and socio-environmental aspects due to higher levels of uncertainty related to them, such as expropriation processes and mitigating or compensatory measures for environmental damage that may be necessary for the implementation of a given project. During the execution of these infrastructure projects, as a rule, the following is noted: i) substantial increases in costs initially foreseen; ii) major delays in implementation schedules; and iii) reductions in originally estimated benefits.

Generally, these situations, when extreme, are the main causes for the construction of works that remain unfinished or for the serious compromise of the viability of the initiatives, making them, in many cases, unfeasible or less feasible than other alternatives that were not prioritized.

As such, GIZ (under the technical coordination of the Federal Court of Accounts of Brazil – TCU) has contracted Oxford Global Projects to undertake the consultancy assignment, "Technical consultancy for conducting a study on the application of reference class forecasting method in public works of the Brazilian transport sector". The objective of the consultancy is to develop a Practical Report with parameters for the application of the Reference Class Forecasting (RCF) method for Brazilian road and railway works (including related fixed links), which, together with other international parameters, can be used as a reference for the development of the method in other OLACEFS member countries in the event of insufficient local data. In so doing, this consultancy is positioned to improve the service offerings of OLACEFS for the active participation of SAIs in national anti-corruption systems, improve the participation of non-governmental actors to increase the scope of SAI audits, strengthen the exchange of technical cooperation and collaboration of CSOs and other government agencies.

3. METHODOLOGY

3.1 Reference Class Forecast

Common traditional project forecasting methods include three-point estimates, Monte Carlo simulations and Earned Value Management (EVM), once project work has started. The use of these methods has led projects to estimate median (50th percentile) or mode (most frequent) accurately, however they also lead to some projects experiencing large cost overruns and schedule delays. Based on historical and statistical analysis, studies by renowned experts indicate that such shortcomings are mainly due to optimism bias or strategic misrepresentation of data in the initial estimates, which are used to justify, in terms of feasibility, a particular infrastructure project.

More specifically, strategic misrepresentation, or political bias, is intentional, purposeful manipulation and/or misstatement of information. Political bias results in artificially low cost and schedule forecasts and benefit overestimates, leading to cost and schedule overruns and benefit shortfalls. As political-organizational pressures increase, project outcomes will increasingly be explained by political bias. It should be noted that the strategic misrepresentation of data constitutes fraud that is usually associated with major corruption schemes, money laundering, and bribery payments to senior public officials responsible for deciding to implement the projects. Therefore, the adoption of internationally recognized mechanisms to reduce the risk of these undesirable factors in feasibility studies of infrastructure projects is considered an extremely important measure to prevent and combat corruption.

Optimism bias is the unintentional tendency to be overly optimistic about future actions, resulting in underestimation of cost and schedule. Due to optimism bias project owners may be ignorant of or underestimate the risk/uncertainties in estimates. Optimism bias is the result of taking an 'inside view', focusing on the project at hand and estimating costs and duration of activities bottom-up.

Instead, RCF is an established estimating approach that deals with political bias and optimism bias by taking an 'outside view' in determining the contingency amount which is based on statistical modelling of outcomes of similar projects. RCF is undertaken in three steps: i) defining the reference class of similar projects and collecting the necessary data, ii) establishing the cumulative probability distribution for the outcomes, iii) making a forecast which dictates the level of certainty of the forecasted value and the corresponding uplift to be added to the bottom-up estimate. Bottom-up estimating works by gathering all the details of a project at the most minute level, whereas in top-down estimating, project managers estimate the project based on previous work on the same or similar projects. Currently, OGP's international database is best suited for bottom-up forecasting, however OGP is developing specific reference classes for RCF to be used for future top-down risk forecasting.

Since RCF uses historical project data as a predictor of the uncertainty and risk of future projects, the effectiveness of RCF depends on the similarity of the reference class. If the project fits well into the reference class, the resulting uplift from the RCF will provide a more reliable estimate of the cost of the project (Awojobi and Jenkins, 2016; Batselier and Vanhoucke, 2016). Moreover, the effectiveness of RCF is influenced by the size of the projects and the size of the reference class (Batselier and Vanhoucke, 2016; Walczak and Majchrzak, 2018); projects need to be sufficiently large, and the reference class should include enough projects. Only if these criteria (similarity, project size, reference class size) are met will RCF outperform other methods. In practical terms, any data is better than no data and a reference class comprising 20-30 past, similar projects is robust enough to derive meaningful insights. Moreover, once data are pooled, they can be analysed to statistically test for similarities between subtypes of projects in the reference class or other characteristics, e.g. size, cost, timelines, location, which might show statistically significantly different risk profiles.

In the final step of RCF, making a forecast, important considerations include risk appetite and evidence-based awareness of how projects within the reference class compare to the project in question. For instance, if the projects in the reference class are generally executed by teams which are more or less experienced than the team of the project in question, this consideration should influence the forecast certainty level selected.

3.2 Inflation adjustment and currency conversion

In the Brazilian data, project costs were stated in nominal (non-inflation-adjusted) national currencies. It was therefore necessary to convert nominal costs into real (inflation-adjusted) costs and then convert them into a common currency. For this work, all costs in international data were converted to annualized 2021 Brazilian real terms. First, price levels were adjusted to 2021 prices using country-specific implicit GDP deflators from the World Bank. Second, all national currencies were converted to Brazilian real terms using purchasing power parity (PPP) exchange rates from the World Bank[4]. They are calculated as an annual average based on monthly averages (local currency units relative to the U.S. dollar).

4 International Comparison Program, World Bank | World Development Indicators database, World Bank | Eurostat-OECD PPP Programme. (n.d.). PPP conversion factor, GDP (LCU per international $). Retrieved March 23, 2023, from 'https://data.worldbank.org/indicator/PA.NUS.PPP'

3.3 Statistical Analysis

We performed the following statistical analysis within each project category: first, basic descriptive statistics were calculated and reported for each data sample. These include the number of projects in each project category as well as the mean, and selected percentiles of cost overruns, schedule overruns and unit costs.

Next, we compared the different metrics by geography in order to be able to make inferences about project performance in Brazil vs project performance in other regions of the world. Whereas descriptive statistics provide a general overview of patterns in the data, including any differences in means or differences in distributions, these differences are not always statistically significant, due to small sample sizes for example. Therefore, another method was used to test whether the Brazilian sample is statistically significantly different from other countries: two-tailed Wilcoxon rank-sum tests, also known as Mann Whitney U tests. The Wilcoxon rank-sum test is used to test whether two samples are likely to derive from the same population (i.e., that the two populations have similarly shaped distributions). This test is sometimes interpreted as a test of the null hypothesis, indicating if the medians of two distributions are equal. The tests were adjusted using Holm-adjustments to control for family-wise error rates. Wilcoxon rank-sum tests are preferable to classic t-tests when the data do not follow normal distributions. For each test completed, median and p-statistics are reported.

4. DATA

This statistical report draws on two different data sources. The cost and schedule data of road and rail projects in Brazil were provided by the Federal Court of Accounts of Brazil (TCU), which is the national agency responsible for auditing public spending. The Brazilian data are divided into two samples: rail and roads/fixed links. The rail sample comprises 39 contracts from three overall rail projects, while the road and fixed link sample comprises 358 contracts. However, the total number of individual projects in the road and fixed link sample could not be reliably determined, as explained below.

TCU is not responsible for managing infrastructure data in Brazil, it only requested access to data made available by government agencies in charge of public works in each sector. Road and fixed link data were obtained from the National Department of Transport Infrastructure (DNIT) through the SIMDNIT[5] system and through the DNIT website[6], accessed in April 2022. Railway data were obtained from the public company Valec (currently Infra S/A) through response to requisition letters[7] in May 2022.

5 http://servicos.dnit.gov.br/simdnit/asp/Main.aspx'
6 'https://www1.dnit.gov.br/editais/consulta/editais2.asp'
7 TC 003.185/2011-7, Letter 495/2022/ADMIN-VALEC/PRESI-VALEC

In addition to the Brazilian data, the cost and schedule data of international projects in the domains of roadways, railways, bridges, and tunnels were obtained from the database of the Oxford Global Projects (OGP). The OGP database consists entirely of project-level data, while the Brazilian datasets are collated at the contract level that may cover only part of the project. This is because it is common for public works contracts in Brazil to end without proper completion of the project, often requiring additional contracts to be signed. However, the data collected in Brazil does not reveal a clear connection between the contracts and the project. As a result, there is a conceptual difference between the data points of the Brazilian databases and the OGP database, and a direct comparison is not initially possible.

To enable a meaningful comparison, the contracts in the Brazilian railway sample were collated into projects by location. However, the statistical comparison is limited by the fact that this resulted in only three total data points. Therefore, any generalizations about the Brazilian railway projects based on the Brazilian data must be made with caution. Nonetheless, the comparison between the Brazilian projects and the international data sheds light on the extent of cost and schedule overruns in Brazil and how they compare with other countries.

4.1 Introduction to the data

Brazilian – Roads and fixed links sample

The Brazilian roads and fixed links dataset contains contracts on roadway, bridges and tunnel constructions. Within the dataset, the variables of interest are '*Object of Intervention*', '*Stretch*', '*Country State*', '*Extension (km)*', '*Contract status*', '*Signature date*', '*Contracted value*', '*Value initially contracted + amendments*', '*Estimated value at the bidding stage*', '*Estimated deadline in the contract signature phase*', and '*Effective deadline or updated forecast*'.

'*Object of Intervention*' refers to the type of road construction. A road construction project can be either a motorway, a bridge, or a tunnel. The '*Stretch*' is the name of the affected road, and '*Country State*' is its location within Brazil. '*Extension (km)*' indicates the length in km of the road project/contract. '*Contract status*' is the status of the contract. A road construction contract can be suspended, in progress or completed. '*Signature date*' indicates the project/contract start date. '*Contracted value*' indicates the estimated project/contract cost (adjusted to BRL 2021). '*Value initially contracted + amendments*', reflects the cost estimate value of the contract/project after the contractor has requested cost and/or schedule estimate modifications (adjusted to BRL 2021). The '*Estimated deadline in the contract signature phase*' reflects the estimated project deadline. '*Effective deadline or updated forecast reflects*' the actual project/contract deadline.

Brazilian – Rails sample

The Brazilian rail dataset contains contracts on railway constructions. Within the dataset, the variables of interest are '*Location of the enterprise*', '*Lot/part*', '*Extension (km)*', '*Contract status*', '*Signature date*', '*Contracted value*', '*Estimated value in the feasibility study phase*', '*Estimated value at the bidding stage*', '*Financial execution percentage of the contract*', '*Value initially contracted + amendments*', '*Estimated deadline in the contract signature phase*', and '*Effective deadline or updated forecast*'. '*Location of the enterprise*' refers to the trainline locations of the project, of which there are three within the dataset, FIOL (West-East Railway), FNS (North-South Railway) and the Southern Extension. '*Lot/part*' contained information important only for the purposes of matching contracts to their respective projects/contracts. '*Extension (km)*' indicates the length in km of the rail contract. '*Contract status*' is the status of the contract. Within the Brazilian rail data, contract statuses were closed, in the process of closing, active or inactive. '*Signature date*' indicates the contract start date. '*Contracted value*' indicates the initial agreed-upon contract cost between the government and the winning bidder (adjusted to BRL 2021). '*Estimated value in the feasibility study phase*' reflects the estimated cost of the contract/project at the time of the feasibility study. '*Estimated value at the bidding stage*' reflects the estimated cost of the contract/project at the time of the bidding stage (adjusted to BRL 2021). '*Contracted value + amendments*' reflects the cost estimate value of the contract/project after the contractor has requested cost and/or schedule estimate modifications (adjusted to BRL 2021). '*Effective payment*' reflects the actual contract/project cost (adjusted to BRL 2021). The '*Estimated deadline in the contract signature phase*' reflects the estimated project deadline. '*Effective deadline or updated forecast*' reflects the actual contract deadline.

Table 1 shows the number of contracts and projects of each project category of the Brazilian data.

CATEGORY	NUMBER OF CONTRACTS	NUMBER OF COMPLETED CONTRACTS	NUMBER OF PROJECTS
Rail	39	18	3
Road	270	193	n/a
Bridge	83	64	n/a
Tunnel	5	5	n/a

Table 1: Summary of Brazilian data

OGP International database

The OGP international database projects on roadway, railway, bridges and tunnel constructions. Within the OGP dataset, the variables of interest included '*Project type*', '*Region*', '*Extension (km)*', '*Signature date*', '*Contracted value*', '*Effective payment*', '*Estimated deadline, and 'Effective deadline*'. '*Project type*' refers to the type of construction. A construction project can be either rail, road, bridge, or tunnel. '*Region*' is the continent on which the project is located. It can be in Europe, South America, North America, Africa, Asia or Oceania. '*Extension (km)*' indicates the length of the construction project in kilometers. Signature date indicates the project start date. '*Contracted value*' indicates the estimated project cost (adjusted to BRL 2021). '*Effective payment*' reflects the actual project cost (adjusted to BRL 2021). The '*Estimated deadline*' reflects the estimated project deadline in the contract signature phase. 'Effective deadline 'reflects the actual project completion date.

The number of projects in each project category as well as some examples of project subtypes that are included in the category are displayed in Table 2.

CATEGORY	NUMBER OF PROJECTS	EXAMPLE OF PROJECT SUBTYPES
Rail	1269	Light rail, conventional rail, urban rail, high-speed rail
Road	3190	Trunk roads, motorways, highways
Bridge	84	Suspension, cable-stayed, lift
Tunnel	127	Cut-and-cover, fixed link, immersed tube

Table 2: Summary of OGP data.

4.2 Challenges

Project benchmarking requires a like-for-like comparison of both delivery costs and outputs. Additionally, calculating unit costs and converting costs into common prices entails selecting and comparing projects that are similar in terms of scope and design.

One problem that was encountered in the course of this analysis is the lack of high-level project characteristics in the Brazilian data (e.g. number of lanes/tracks, urban/rural construction, greenfield/brownfield construction side, etc.). As projects usually vary in terms of context or design, it was not possible to meaningfully compare Brazilian unit costs with international data.

Another issue affecting the implications of this analysis is the different data foundation between the datasets. The international data provided by OGP is project-based, in contrast to the Brazilian datasets, which consist of contracts that are not necessarily individual projects. Since the effectiveness of RCF depends upon the similarity of the reference class, OGP first had to identify the contracts which belonged to each individual project in order to compare Brazilian data to the international datasets.

According to Law No 8,666/1993, 10,520/2002, 12,462/2011 and 14,133/2021, which is currently the main legal framework for the procurement of government contracts in Brazil, if works, services, or purchases contracts exceed their intended cost by 25% and renovation/reform of building or equipment contracts exceed their intended cost by 50%, a new contract may be written (through a new bidding process) to continue the project. The new contract does not have to reflect or reference the originally intended cost or duration of the project at hand, thereby rendering accurate calculations of cost and schedule overruns of projects, which consist of multiple contracts, highly challenging. This legal circumstance can be seen clearly in the Brazilian datasets. Figure 1 and Figure 2 show that the data are truncated at 25% cost overrun in both Brazilian data samples. Within the Brazilian road and fixed links data, all contracts appear to experience cost overruns of 25% or lower (Figure 1).

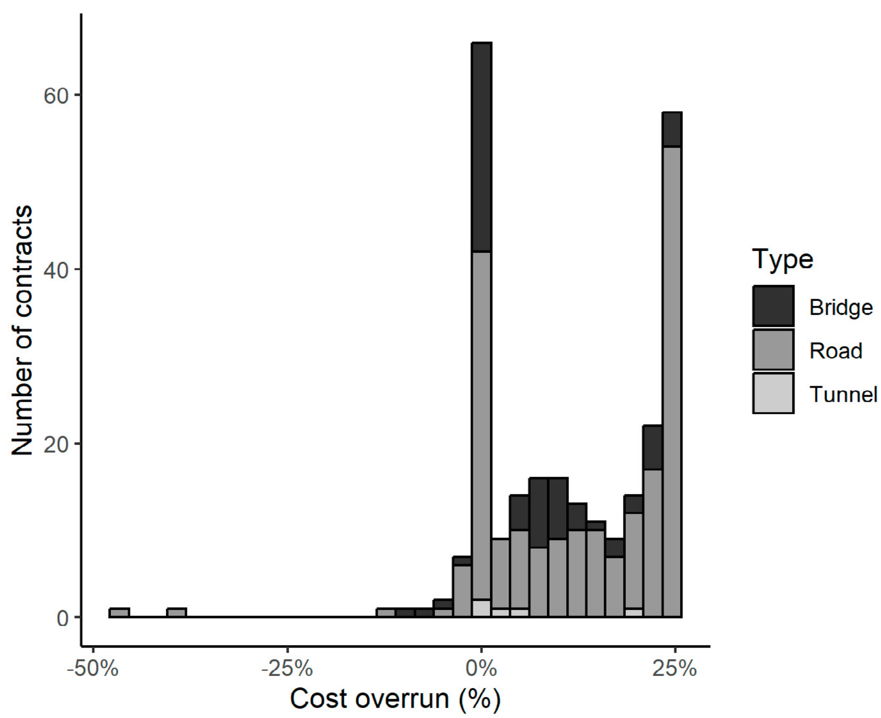


Figure 1: histogram displays the cost overrun values of completed contracts and their frequencies across the Brazilian road and fixed links dataset.

A similar trend is apparent in the Brazilian rail dataset (Figure 2). We can observe a high accumulation of contracts at cost overrun = 25%.

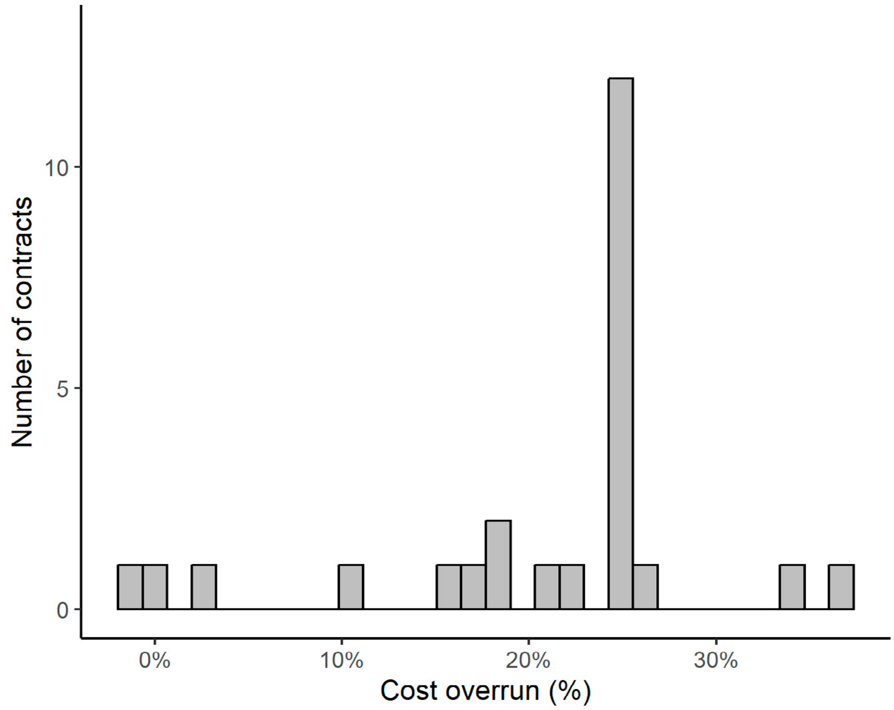


Figure 2: histogram displays the cost overrun values of completed contracts and their frequencies across the Brazilian rail dataset.

For this reason, TCU endeavored to update the datasets to reflect whole projects instead of contracts, which resulted in decreased sample size for the rail sample as contracts were grouped into their appropriate projects. However, even after this adjustment Brazilian data still show a spike at 25% overruns, which suggests that collating contracts into same work packages was not completely successful. The data most likely still represent a mixture of contracts and projects.

Considering this problem, an additional approach was proposed for the Brazilian rail sample, which allowed to compile contract data into projects by assigning each of them to one the three major rail lines in Brazil, FIOL (West-East Railway), FNS (North-South Railway) and the Southern Extension. OGP used the three major Brazilian rail lines as proxies for three individual projects. Unfortunately, no such alternative way of grouping contracts together was feasible for the Brazilian road and fixed links sample.

4.3 Calculation of performance metrics

For the calculation of all performance metrics only projects with status "completed" have been considered.

Cost overrun is calculated as:

$$\frac{\textit{Effective payment}}{\textit{Contracted value} \cdot \textit{Financial execution percentage}}$$

where the upper term is the *actual project cost in 2021 BRL* and the lower term is the *estimated project cost in 2021 BRL* multiplied by the *project complete percentage*.

Schedule overrun is calculated as:

$$\frac{\textit{Effective deadline or updated forecase} - \textit{Signature date}}{\textit{Estimated deadline in the contract} - \textit{Signature date} \cdot \textit{Financial execution percentage}}$$

where the upper term is *actual project duration* in days and the lower term is *estimated project duration multiplied* by the *project complete percentage*.

Unit cost, or cost per lane-kilometer, is calculated as:

$$\frac{\textit{Effective payment}}{\textit{Extension (km)} * \textit{Number of lanes}}$$

where the upper term is the is the *actual project cost in BRL* and the lower term measures project size in lane-km (accordingly track-km for rail projects).

Notice that for the calculation of Cost and Schedule Overruns, the '*percentage of financial execution*' is used as a correction term. This is because projects from the Brazilian dataset can be finished (project status = '*completed*'), but incomplete in terms of the original scope of the contract. Hence, cost and duration do not reflect the originally projected work, but only a part of it. To take this into account, hypothetical cost and duration at completion is estimated for incompletely finished projects. This is done by dividing effective cost and duration by the percentage of project completion. Notice further that a linear type of project development is assumed. Meaning that if, for example, 20% of the project's scope is missing, it will be estimated that finishing the contract takes 20% more extra time and money. Furthermore, '*Financial execution percentage*' is used as a proxy for physical project completion, since there was no complete data available on actual physical project completion status in the Brazilian dataset.[8]

8 For the available data it was found that the financial progress of Brazilian projects, on average, is about 1% off the physical progress. Financial execution percentage therefore seems to be a valid proxy for actual physical execution percentage.

The ’Financial execution percentage of the contract’ is calculated as follows:

$$\frac{\textit{Contracted value} + \textit{amendments}}{\textit{Effective payment}}$$

in other words, the percentage of the ‘Contracted value + amendments’ that has been covered by the ‘Effective payment’.

Notice that with this definition of ’Financial execution percentage of the contract’, the calculation of cost overruns essentially boils down to the following:

$$\frac{\textit{Contracted value} + \textit{amendments}}{\textit{Contracted value}}$$

meaning that the cost overrun is measured by the amendments added to the contract, i.e., the cost estimate modifications requested by the contractor.

The OGP data, on the other hand, only include completed projects. Hence a correction as explained above is not necessary for other international projects.

In addition to the previous described calculation for cost overruns using the ‘*Contracted value*’ as estimated project cost, further measures have been considered within the Brazilian data. These additional cost overruns are calculated with the ‘*Estimated value at the feasibility stage*’ and the ‘*Estimated value at the bidding stage*’ as measures for estimated project costs. Of the three cost overrun calculations, cost overruns calculated from feasibility stage estimates have the greatest variability across the Brazilian dataset and these cost overruns do not appear to experience the same cost overrun truncation at 25% which is observed in the other two cost overrun calculations. Ideally the ‘*Estimated value at the feasibility*’ stage should be used as measure for estimated project cost in this analysis. However, this was not feasible due to the small number of observations containing data on this variable. On the other side, when varying between contracted ‘*Contracted value*’ and ‘*Estimated value at the bidding stage*’, no significant differences in the resulting cost overruns were found.

After identifying which contracts formed entire projects, OGP calculated the project-wide cost overruns as follows. The ‘*Contracted value*’ for the whole project is kept equal to that of the first contract of the project. ‘*Value initially contracted + amendments*’ is calculated as the sum of the actual costs of every contract within the project.

Similarly, the schedule overruns are calculated as follows. *Estimated project duration* is the estimated duration of the first contract and the *actual project duration* is calculated as the sum of the actual durations of each contract of the project.

For the additional analysis on Brazilian railways, in which all projects are collapsed into three projects corresponding to rail line location, FIOL (West-East Railway), FNS (North-South Railway) and the Southern Extension, calculations were conducted as follows. The *Contracted value* for each rail line location was calculated as the sum of the estimated costs of every project, with supplementary contracts excluded from the calculation. The *Effective payment* for each rail line location was calculated as the sum of the actual costs of every project in each location, with supplementary contracts included in the calculation. Cost overruns are then calculated as indicated above. The estimated project duration for each rail line location was calculated as the sum of the estimated durations for every project in each location and the actual project duration is calculated as the sum of the actual durations of every project. Schedule overruns are then calculated as explained above.

5. ANALYSIS

In the presentation of the results in the following section, RCF 50 is the median and RCF 80 is the 80th percentile. For instance, if 80% of European road projects in the reference class had a cost overrun of 50% or less compared to the base cost estimate, we express that as follows: *RCF 80 = 50%*. We adopted this language to help projects clearly differentiate between bottom-up risk estimates, which refer to e.g. P50 and P80, and the results of the reference class analyses. In the tables below, the Brazilian rail data sample are presented both on contract level (n=18) and on project level (n=3).

5.1 Cost Overrun

5.1.1 Descriptive statistics

Table 3 **shows a descriptive overview for cost overruns for road, rail, bridge, and tunnel data.** Notice that Brazilian figures stem from the Brazilian dataset, which is measured on the contract level. It stands in contrast to the other international data from the OGP database, which is at the project level. However, there is an additional row for Brazilian rail projects labeled as "Brazil (projects)", which shows statistics for consolidated Brazilian data grouped into projects by location. This consolidation of contract data was not feasible for any other type than rail.

TYPE	REGION	SAMPLE SIZE	MEAN	FREQUENCY OF OVERRUN	MEDIAN (RCF 50)	RCF 80	HISTORICAL RANGE
Road	Brazil	193	12%	8 out of 10	14%	25%	2005 – 2021
	Europe	1645	12%	5 out of 10	1%	21%	1969 – 2016
	South America	41	53%	9 out of 10	59%	59%	1998 – 2007
	Africa	11	67%	8 out of 10	29%	65%	1990 – 2004
	Asia	261	18%	7 out of 10	13%	29%	1982 – 2007
	Oceania	40	43%	7 out of 10	12%	65%	1995 – 2020
	North America	76	16%	5 out of 10	0%	18%	1941 – 2021
	Overall	2267	14%	6 out of 10	4%	24%	1941 – 2021
Rail	Brazil	18	27%	9 out of 10	25%	30%	2002 – 2017
	Brazil (projects)	3	38%	10 out of 10	39%	44%	2002 – 2017
	Europe	238	32%	7 out of 10	12%	55%	1954 – 2017
	South America	2	25%	5 out of 10	25%	45%	1976 – 1985
	Africa	1	71%	10 out of 10	71%	71%	2002
	Asia	67	44%	7 out of 10	20%	66%	1966 – 2011
	Oceania	22	8%	5 out of 10	0%	14%	2013 – 2021
	North America	198	28%	7 out of 10	15%	54%	1898 – 2017
	Overall	546	31%	7 out of 10	14%	55%	1898 – 2021
Bridge	Brazil	64	7%	6 out of 10	6%	16%	2005 – 2019
	Europe	29	27%	6 out of 10	15%	61%	1962 – 2006
	Asia	6	34%	8 out of 10	14%	87%	1985 – 2009
	Oceania	1	113%	10 out of 10	113%	113%	1932
	North America	19	24%	6 out of 10	3%	33%	1869 – 2016
	Overall	119	17%	6 out of 10	7%	23%	1869 – 2016
Tunnel	Brazil	5	5%	6 out of 10	3%	7%	2006 – 2013
	Europe	54	35%	7 out of 10	25%	71%	1963 – 2016
	South America	1	79%	10 out of 10	79%	79%	1939
	Asia	4	14%	10 out of 10	14%	22%	1986 – 2007
	Oceania	7	30%	10 out of 10	33%	40%	1982 – 2010
	North America	4	81%	8 out of 10	57%	128%	1919 – 2007
	Overall	75	33%	7 out of 10	23%	64%	1919 – 2016

Table 3: overview of cost overruns in international data used for reference class construction.
Note: Brazilian data listed here are at the contract level compared to the rest of the data which are at project level. An exception is 'Brazil (projects)' which estimates a project level by consolidating contracts by location.

Figure 3 compares the cost overruns of road, rail, bridge, and tunnel projects by region. The observation of interest is the median cost overrun value within each region, represented by the bold horizontal line in each boxplot. These measures serve as valuable reference points to which similar projects should be compared.

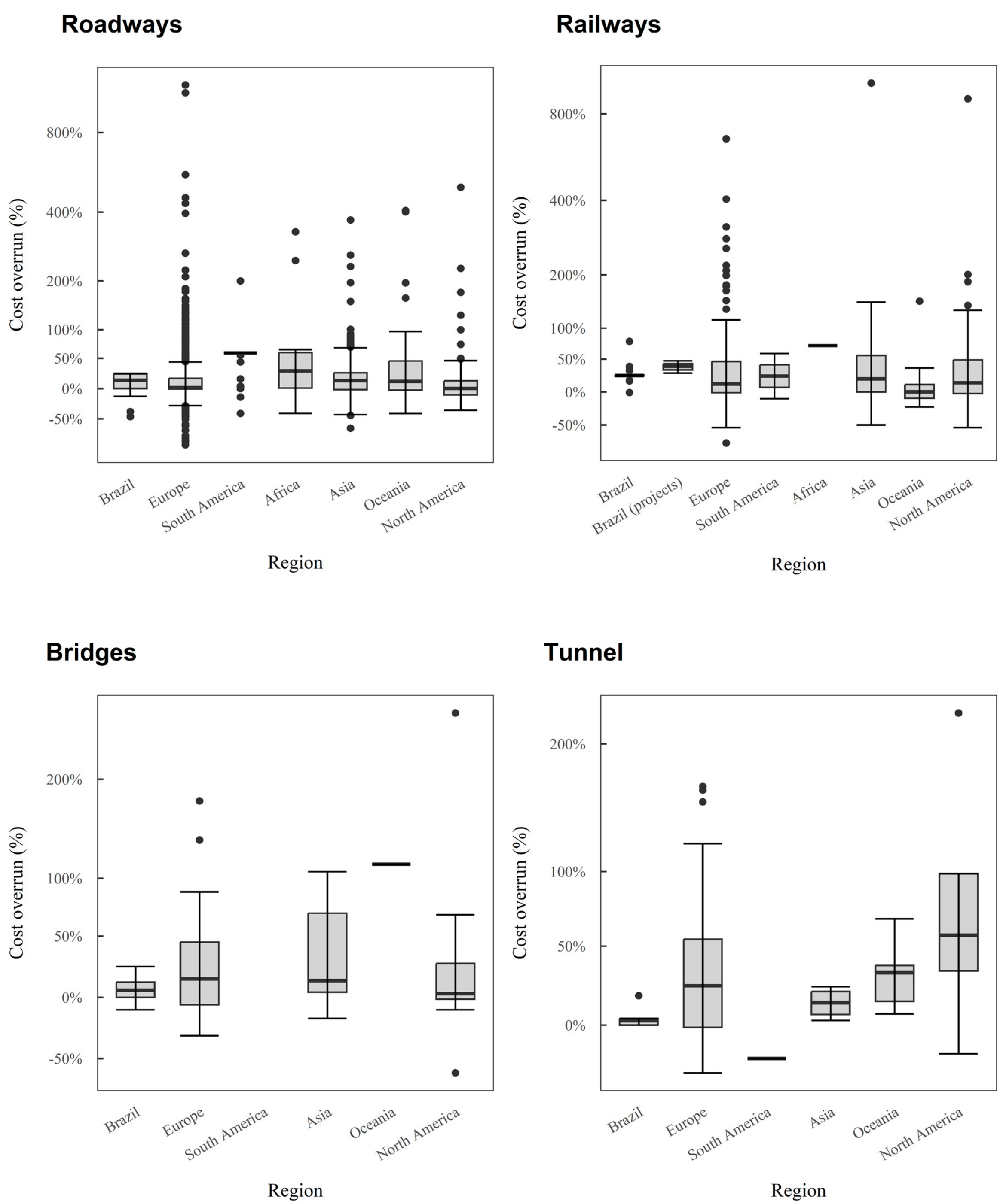


Figure 3: boxplots comparing the cost overrun across relevant project types within the OGP database.

5.1.2 Statistical Tests

Table 4 reports the results of Wilcoxon rank-sum tests that are used to identify whether there is a statistically significant difference in the distribution of cost overrun in Brazil versus in other country groups. The Wilcoxon rank-sum tests in this study indicate that the cost overruns of **road projects** are higher in *Brazil (Mdn = 14%)* when compared to *Europe (Mdn = 1%)* and *North America (Mdn = 0%)* but lower than for *South America (Mdn = 59%)*. Furthermore, statistically significant differences in **rail projects** have been found between *Brazil (Mdn = 25%)* and *Oceania (Mdn = 0%)*. **For bridge and tunnel project types** there are no statistically significant differences with Brazil.

DIFFERENCES BETWEEN BRAZIL COST OVERRUN AND	PROJECT TYPE	p-VALUE
Europe	Road	5e-09***
North America	Road	2e-05***
South America	Road	2e-13***
Oceania	Rail	3e-04***

Table 4: Wilcoxon rank-sum tests for statistical significance of differences in cost overruns

Note: Statistical significance indicators: Reject null hypothesis that samples derive from the same distribution at the following levels: * p<5%; ** p<1%; ***p<0.1%

5.2 Schedule Overrun

5.2.1 Descriptive statistics

Table 5 shows a descriptive overview for schedule overruns for road, rail, bridge, and tunnel data. Notice that Brazilian figures stem from the Brazilian dataset, which is measured on the contract level. It stands in contrast to the other international data from the OGP database, which is at the project level. However, there is an additional row for Brazilian rail projects labeled as "Brazil (projects)", which shows statistics for consolidated Brazilian data grouped into projects by location. This consolidation of contract data was not feasible for any other type than rail.

TYPE	REGION	SAMPLE SIZE	MEAN	FREQUENCY OF OVERRUN	MEDIAN (RCF 50)	RCF 80	HISTORICAL RANGE
Road	Brazil	193	103%	10 out of 10	76%	147%	2005 – 2021
	Europe	249	39%	6 out of 10	15%	64%	1971 – 2016
	South America	15	34%	6 out of 10	17%	70%	1998 – 2017
	Africa	6	133%	10 out of 10	140%	209%	1992 – 2003
	Asia	127	28%	6 out of 10	14%	50%	1985 – 2007
	Oceania	37	1%	4 out of 10	0%	9%	1989 – 2011
	North America	23	39%	4 out of 10	0%	21%	1941 – 2021
	Overall	650	55%	7 out of 10	27%	100%	1941 – 2021
Rail	Brazil	18	405%	10 out of 10	407%	497%	2002 – 2017
	Brazil (projects)	3	439%	10 out of 10	445%	457%	2002 – 2017
	Europe	24	45%	8 out of 10	18%	85%	1974 – 2011
	Asia	46	19%	5 out of 10	8%	50%	1971 – 2011
	Oceania	12	10%	5 out of 10	2%	26%	Unavailable
	North America	50	40%	7 out of 10	20%	55%	1898 – 2016
	Overall	150	76%	7 out of 10	20%	100%	1898 – 2016
Bridge	Brazil	64	145%	10 out of 10	100%	187%	2005 – 2019
	Europe	15	22%	7 out of 10	19%	37%	1967 – 2010
	Asia	6	26%	7 out of 10	20%	76%	1989 – 2009
	North America	4	10%	5 out of 10	7.5%	24%	1927 – 1997
	Overall	89	110%	9 out of 10	4163	142%	1927 – 2010
Tunnel	Brazil	5	25%	6 out of 10	0%	24%	2006 – 2013
	Europe	16	23%	6 out of 10	2%	36%	1976 – 2016
	Asia	2	-1%	5 out of 10	-1%	7%	1986 – 1997
	Oceania	7	6%	4 out of 10	-1%	9%	1982 – 2007
	North America	2	85%	10 out of 10	85%	86%	1987 – 2007
	Overall	32	22%	6 out of 10	1%	42%	1976 – 2016

Table 5: overview of schedule overruns in international data used for reference class construction.
Note: Brazilian data listed here is at the contract level compared to the rest of the data which is at project level. An exception is 'Brazil (projects)' which estimates a project level by consolidating contracts by location.

Figure 4 compares the schedule overruns of road, rail, bridge and tunnel projects by region. The observation of interest is the median cost overrun value within each region, represented by the bold horizontal line in each boxplot. These measures serve as valuable reference points to which similar projects should be compared.

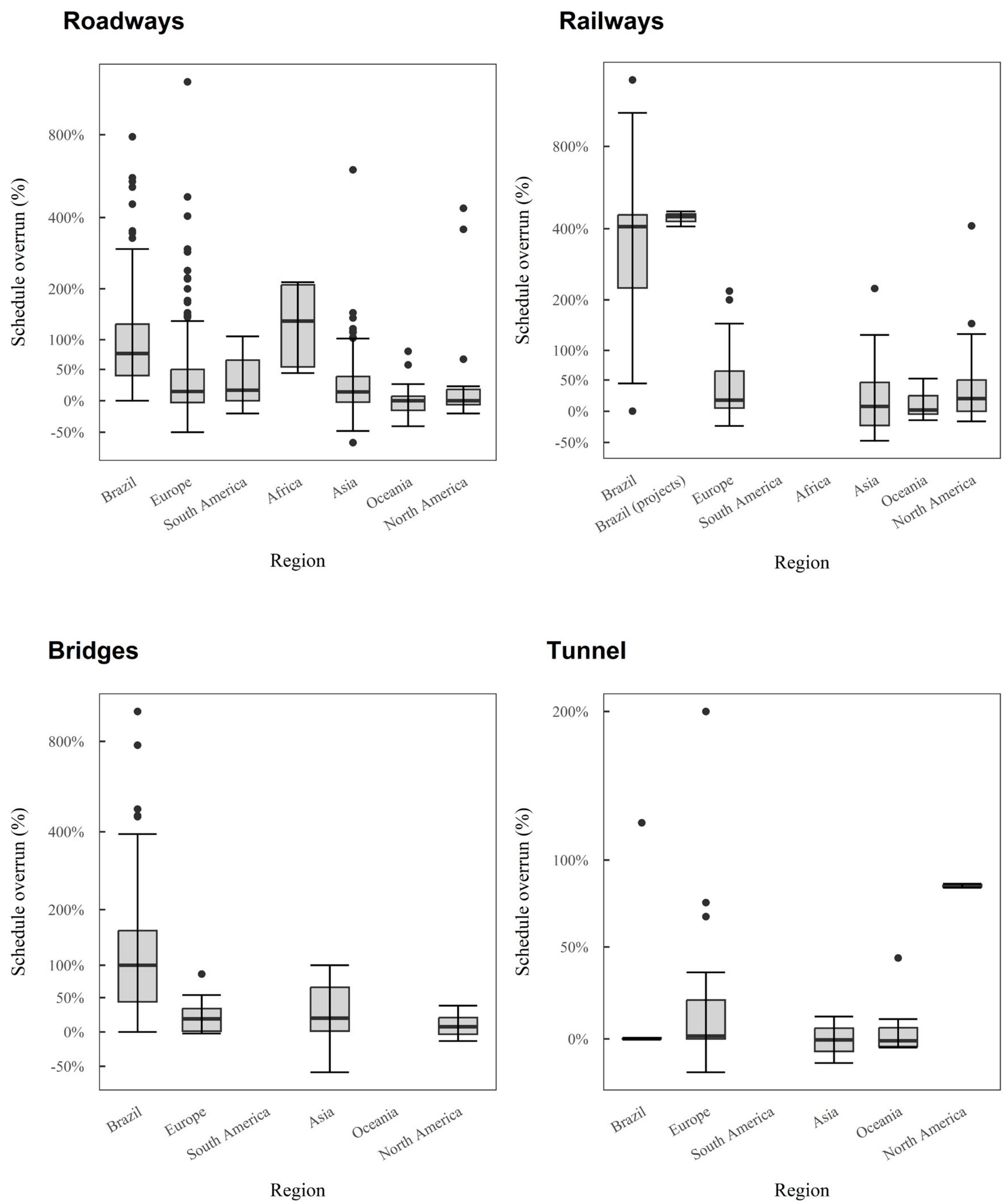


Figure 4: boxplots comparing the schedule overrun across relevant project types within the OGP database.

5.2.2 Statistical Tests

Table 6 reports the results of Wilcoxon rank-sum tests that are used to identify whether there is a statistically significant difference in the distribution of schedule overruns in Brazil versus in other country groups. *Brazil (Mdn = 76%)* experiences greater schedule overruns in **road project** when compared with *Asia (Mdn = 14%)*, *Europe (Mdn = 15%)*, *North America (Mdn = 0%)*, *Oceania (Mdn = 0%)* and *South America (Mdn = 17%)*. **Rail projects** have also greater schedule overruns in *Brazil (Mdn = 407%)* when compared to *Asia (Mdn = 8%)*, *Europe (Mdn = 18%)*, *North America (Mdn = 20%)* and *Oceania (Mdn = 2%)*. Furthermore, statistically significant differences in schedule overruns for **Bridge projects** have been found. It seems that bridge construction projects in *Brazil (Mdn = 100%)* have greater schedule overruns than in *Europe (Mdn = 19%)*. There is also some weak evidence *(p-value < 6%)* that schedule overruns for bridge projects is greater compared to *North America (Mdn = 8%)* For **tunnel projects** there are no statistically significant differences with Brazil.

DIFFERENCES BETWEEN BRAZIL SCHEDULE OVERRUN AND	PROJECT TYPE	p-VALUE
Asia	Road	2e-16***
Europe	Road	2e-16***
North America	Road	5e-07***
Oceania	Road	4e-16***
South America	Road	0.013*
Asia	Rail	4e-08***
Europe	Rail	7e-06***
North America	Rail	1e-07***
Oceania	Rail	2e-06***
Europe	Bridge	7e-04***
North America	Bridge	0.055

Note: Statistical significance indicators: Reject null hypothesis that samples derive from the same distribution at the following levels:
* p<5%; ** p<1%; ***p<0.1%

Table 6: Wilcoxon rank-sum tests for statistical significance of differences in schedule overruns

5.3 Cost per lane kilometer

5.3.1 Descriptive statistics

Table 7 **shows a descriptive overview for cost per kilometer (BRL, 2021) for international road, rail, bridge, and tunnel data.** It was not possible to calculate the cost per lane kilometer in the Brazilian data since no information was available on actual project size, such as the number of lanes (tracks for rail respectively) or road width. Furthermore, the contracts within the Brazilian datasets do not consider the same costs as the projects within the OGP data. OGP data include design costs, expropriation, supervision, management of environmental impacts, and supply costs. Thus, a comparison of unit costs in absolute values between international and Brazilian data is not feasible.

TYPE	REGION	SAMPLE SIZE	MEAN	MEDIAN (RCF 50)	RCF 80	HISTORICAL RANGE
Road	Europe	182	19	6	11	1969 – 2016
	South America	1	8	8	8	2007
	Asia	53	23	12	20	1982 – 2007
	Oceania	5	16	17	27	1995 – 2020
	North America	12	8	4	11	1941 – 2021
	Overall	253	18	5	11	1941 – 2021
Rail	Europe	11	77	52	115	1954 – 2017
	Asia	12	24	24	34	1966 – 2011
	North America	27	95	52	96	1898 – 2017
	Overall	50	55	27	70	1898 – 2021
Bridge	Europe	8	127	114	201	1962 – 2006
	Asia	5	241	82	330	1985 – 2009
	North America	2	65	65	89	1869 – 2016
	Overall	15	59	5	78	1869 – 2016
Tunnel	Europe	13	478	50	152	1963 – 2016
	South America	1	34	34	34	1939
	Asia	1	15	15	15	1986 – 2007
	Overall	15	324	49	101	1919 – 2016

Table 7: overview of cost per lane km in million BRL; international data used for reference class construction.

Figure 5 compares the cost per lane kilometer of road, rail, bridge, and tunnel projects by region. The observation of interest is the median unit cost value within each region, represented by the bold horizontal line in each boxplot. These measures serve as valuable reference points to which similar projects should be compared.

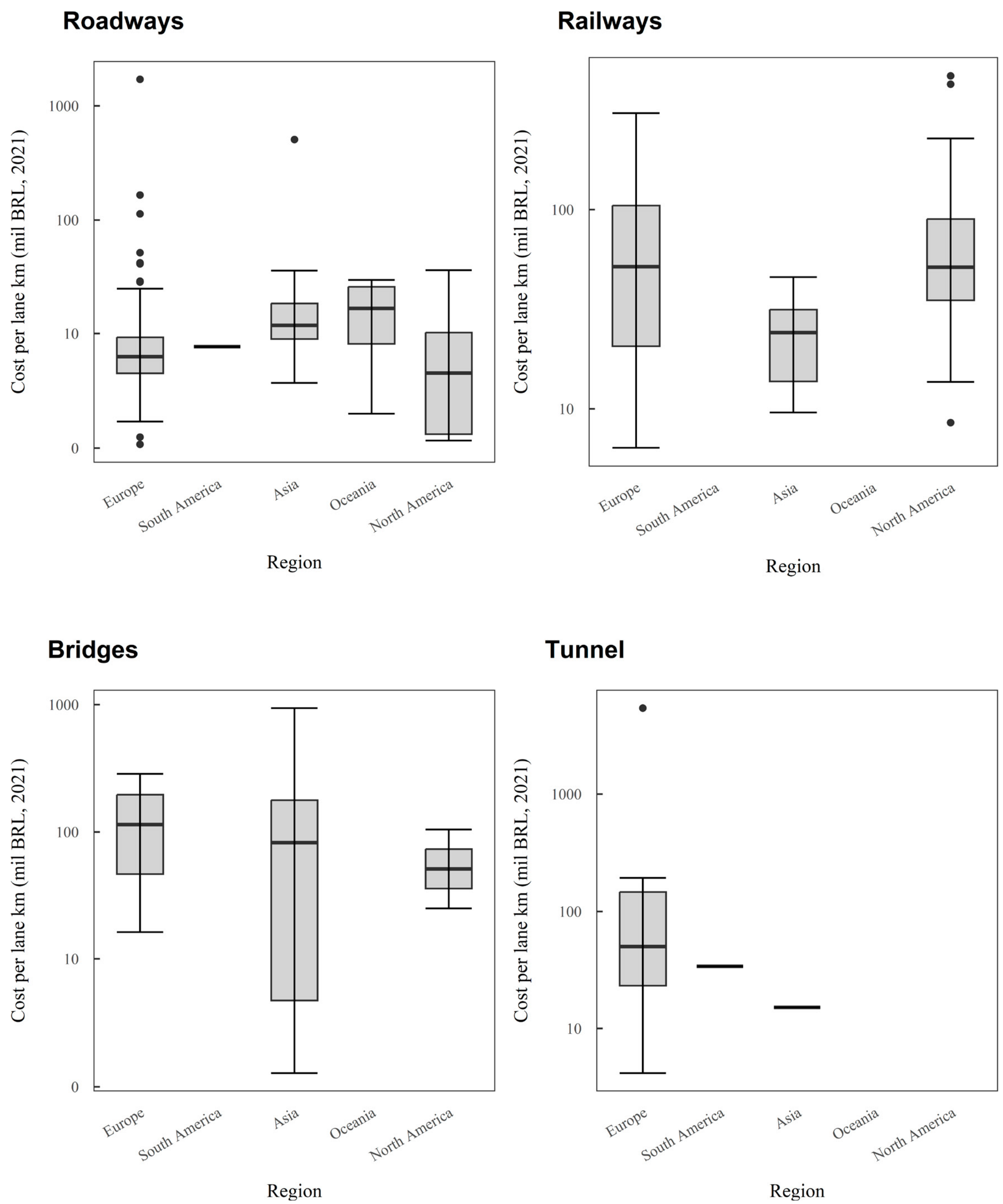


Figure 5: boxplots comparing the cost per kilometer across relevant project types within the OGP database.

5.4 Brazilian railway dataset consolidated

After consolidating the Brazilian railway dataset into projects according to location, OGP extracted the cost and schedule overruns of the following three projects: FIOL (West-East Railway), FNS (North-South Railway) and Southern Extension.

Figure 6 and Figure 7 below demonstrate how the three major rail projects in Brazil compare to international data in terms of cost and schedule overruns. The FIOL (West-East Railway) rail project was 29% over budget, the FNS (North-South Railway) rail project was 47% over budget and the Southern Extension line was 39% over budget. Figure 7 includes color-coded dashed lines for ease of viewing the intersection of each cost overrun value with the RCF curve. In terms of project completion time, the FIOL (West-East Railway) rail project was 465% over schedule, the FNS (North-South Railway) rail project was 445% over schedule and the Southern Extension line was 408% over schedule. As seen in Figure 8, none of the three schedule overruns intersect with the RCF curve since all three schedule overruns fall far above the RCF95 percentile of international data within the OGP database.

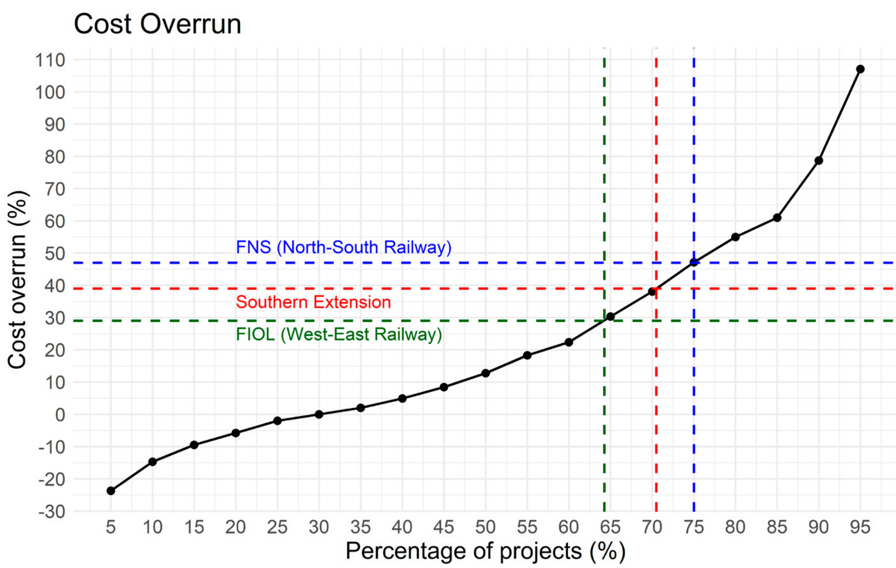


Figure 6: Reference class for cost overrun in international rail projects.

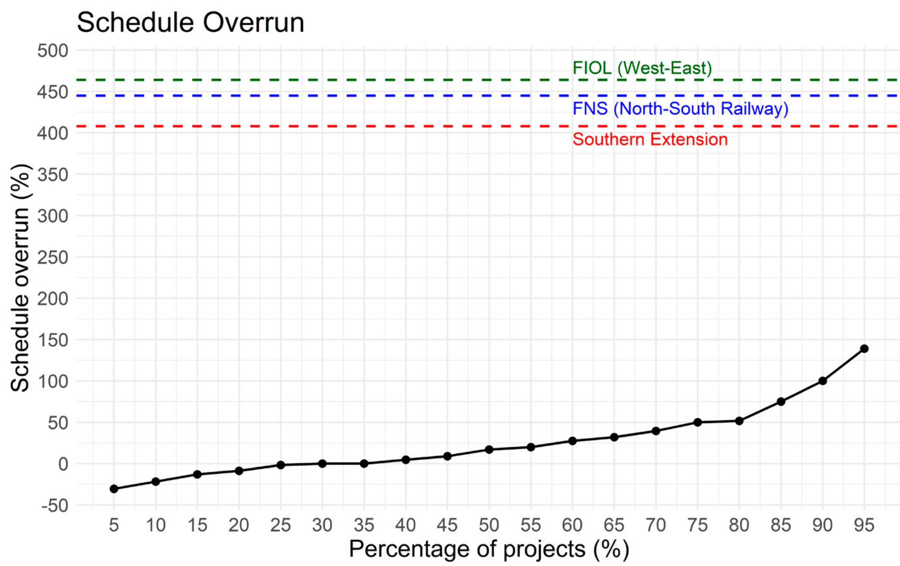


Figure 7: Reference class for schedule overrun in international rail projects.

In other words, the comparisons to international data reveal a discrepancy between cost and schedule overruns. Brazilian data appear to be comparable to the international data regarding cost overruns, however schedule overruns in the Brazilian rail projects far exceed schedule overruns in international data. It is always necessary to point out that cost increases for Brazilian projects may be greater than those calculated here, since there is no information on all the contracts that make up a project.

6. RESULTS

Table 8 **summarizes key conclusions from this analysis, based on differences in median cost and schedule and the statistical tests of differences in the distribution between countries.** We found statistically significant differences between Brazil and some other countries for some project types, but not others.

PROJECT TYPE	COST OVERRUN	SCHEDULE OVERRUN
Motorways	There is **evidence of statistically significant differences** in the cost overruns when building motorways between Brazil and other country groupings. Brazil seems to have ***higher*** cost overruns when compared to Europe and North America. On the other hand, Brazil seems to have ***lower*** cost overruns when compared to *South America*, always noting that Brazilian data are incomplete, and such values may perhaps be higher when considering all costs.	There is **evidence of statistically significant differences** in the schedule overruns when building motorways between Brazil and other country groupings. Brazil seems to have ***higher*** schedule overruns when compared to *Asia*, *Europe*, *North America*, *Oceania* and *South America*.
Railways	There is **little evidence of statistically significant differences** in the cost overruns when building railways between Brazil and other country groupings. Brazil seems to have ***higher*** overruns when compared to *Oceania*. It seems that Brazilian cost overruns in rail projects are generally comparable to those in other regions of the world.	There is **evidence of statistically significant differences** in the schedule overruns when building railways between Brazil and other country groupings. Brazil seems to have ***higher*** schedule overruns when compared to *Asia*, *Europe*, *North America* and *Oceania*. It seems that Brazilian schedule overruns in rail projects are generally greater than those in other regions of the world.
Bridges	There are **no statistically significant differences** in the cost overruns in Brazil versus other country groupings. It seems that Brazilian cost overruns in bridge projects are generally comparable to those in other regions of the world.	There is **some evidence of statistically significant differences** in the schedule overruns when building bridges between Brazil and other country groupings. Brazil seems to have ***higher*** schedule overruns when compared to *Europe* and North America
Tunnel	There are **no statistically significant differences** in the cost overruns when building tunnel in Brazil versus other country groupings. This result might be driven by the small sample sizes in this category.	There are **no statistically significant differences** in the schedule overruns when building tunnel in Brazil versus other country groupings. This result might be driven by the small sample sizes in this category

Nonetheless, this analysis has several important limitations and conclusions should be drawn from it with caution.

First, there are a small number of Brazilian tunnel projects, which means that it is difficult to identify statistically significant differences in costs and schedules or to generalize from these findings to other projects.

Second, information was not available on some project attributes that affect cost, such as a more exact measure for the construction size, other than the extension in km (e.g. number of lanes/tracks, size of the road/railway), urban or rural location, topography, and design specifications in general.

Third, as mentioned previously, the data are structured in a fundamentally different way. The Brazilian data are collected on a contract level, whereas data in the OGP database consist of complete projects that can involve several contracts until completion. Efforts to resolve this problem by aggregating Brazilian contracts into projects have been mostly unsuccessful. The discrepancy between cost and schedule overruns within Brazilian data further illustrates the issue: schedule overruns are far higher than what would be expected given the cost overruns in Brazilian data.

Fourth, and finally, Brazilian legislation dictates that government infrastructure contracts in Brazil must be ended if they exceed their intended cost by 25% and hence a new contract must be written to be able to continue work on the project. This means that the data, as it stands, cannot be compared to other contract data from international projects either.

Consequently, **our recommendation is to use international data for benchmarking and Reference Class Forecasting, until the previously summarized limitations have been addressed**.

For this purpose, Table 9 **and the following Figures 9 to 14 show the distributions for international reference classes derived from the OGP database.** These should serve as the benchmark for infrastructure project outcomes.

	ROADWAYS			RAILWAYS			BRIDGES			TUNNEL		
Percentage of projects (%)	Cost overrun (%)	Schedule overrun (%)	Cost per lane km (mil BRL)	Cost overrun (%)	Schedule overrun (%)	Cost per lane km (mil BRL)	Cost overrun (%)	Schedule overrun (%)	Cost per lane km (mil BRL)	Cost overrun (%)	Schedule overrun (%)	Cost per lane km (mil BRL)
5%	-13%	-21%	-29%	0.9	-24%	-30%	8.0	-17%	-11%	2.3	-22%	3.5
10%	-12%	-12%	-20%	1.9	-15%	-22%	9.8	-12%	-1%	8.3	-17%	5.7
15%	-5%	-7%	-15%	2.6	-9%	-13%	13.0	-10%	0%	16.2	-10%	9.4
20%	-4%	-4%	-9%	3.0	-6%	-9%	15.2	-6%	0%	22.4	-4%	13.1
25%	-3%	-1%	-3%	3.7	-2%	-2%	21.6	-4%	0%	26.8	0%	18.3
30%	0%	0%	0%	4.4	0%	0%	27.8	-2%	0%	34.2	5%	24.5
35%	0%	0%	0%	4.6	2%	0%	30.0	0%	3%	50.5	11%	32.0
40%	0%	0%	0%	5.1	5%	5%	34.1	2%	5%	70.0	18%	33.2
45%	1%	1%	7%	5.6	8%	9%	38.5	4%	13%	88.1	20%	37.2
50%	1%	4%	11%	6.4	13%	17%	42.7	9%	16%	103.8	24%	46.5
55%	3%	6%	15%	6.7	18%	20%	46.1	18%	19%	109.9	26%	48.5
60%	9%	9%	21%	7.4	22%	27%	49.6	22%	22%	113.2	30%	52.4
65%	11%	13%	31%	8.4	30%	32%	52.2	24%	29%	120.5	33%	60.2
70%	13%	16%	34%	9.6	38%	39%	61.8	32%	35%	164.3	43%	82.4
75%	26%	20%	48%	10.8	47%	50%	77.5	42%	35%	184.4	52%	117.2
80%	42%	24%	60%	12.1	55%	52%	93.2	62%	39%	194.7	68%	147.3
85%	68%	30%	80%	14.7	61%	75%	107.7	69%	48%	203.7	72%	153.1
90%	78%	46%	102%	20.2	79%	100%	160.4	88%	67%	253.3	108%	177.0
95%	85%	64%	140%	25.0	107%	139%	268.5	120%	84%	481.0	157%	1754.6
N	32	2267	650	253	546	150	50	119	89	15	75	15
Average	16%	10%	25%	8.1	23%	26%	60.9	23%	21%	118.3	32%	148.2

Table 9: Distributions for cost overrun, schedule overrun and cost per km from international reference classes

6.1 Cost Overrun

The RCF curves in Figure 8 and Figure 9 show the cumulative probability distribution of cost overrun values within the P5 – P95 percentile interval in international data. Notice that 50% of the roadway projects had a cost overrun of 4% or less, and 80% of the roadway projects had a cost overrun of 24% or less. Furthermore, 50% of the railway projects had a cost overrun of 13% or less and 80% of the railway projects had a cost overrun of 55% or less.

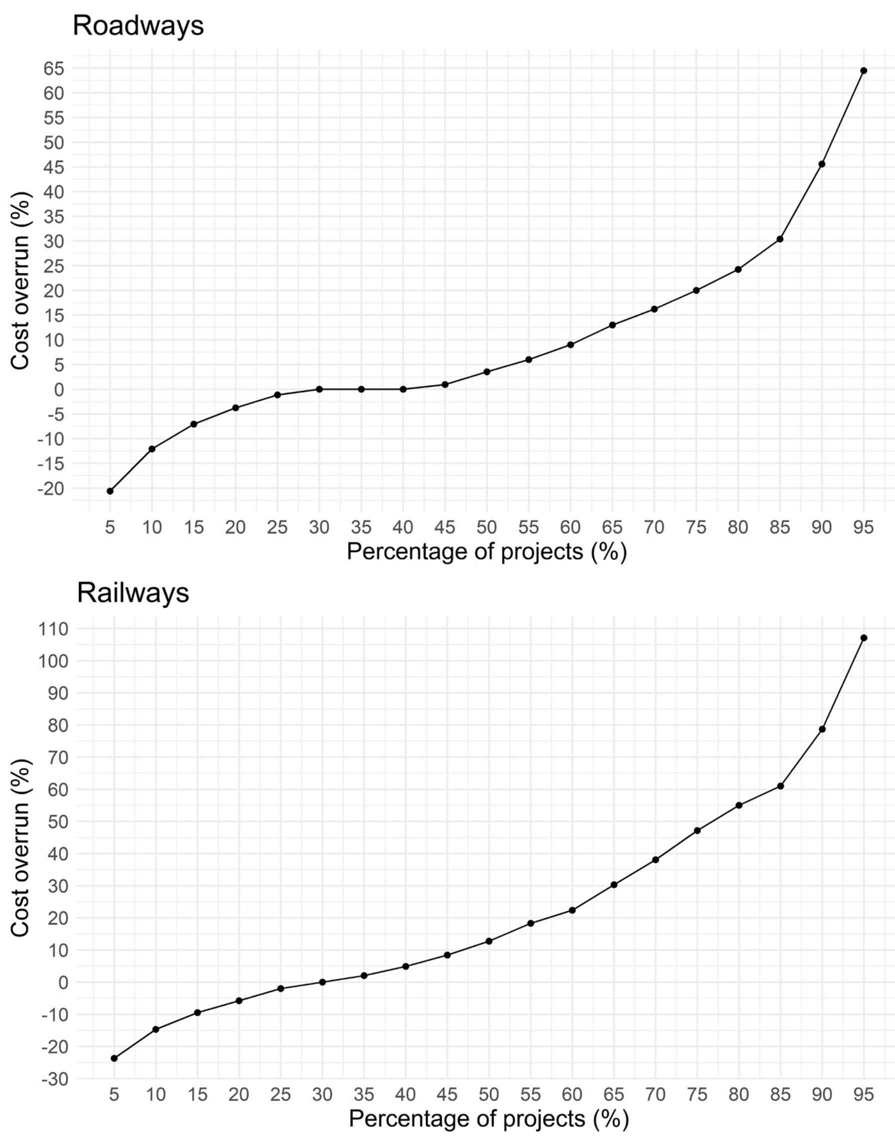


Figure 8: RCF of cost overruns in international roadway and railway projects

50% of the bridge projects had a cost overrun of 9% or less, and 80% of the bridge projects had a cost overrun of 62% or less. Furthermore, 50% of the tunnel projects had a cost overrun of 24% or less and 80% of the tunnel projects had a cost overrun of 68% or less.

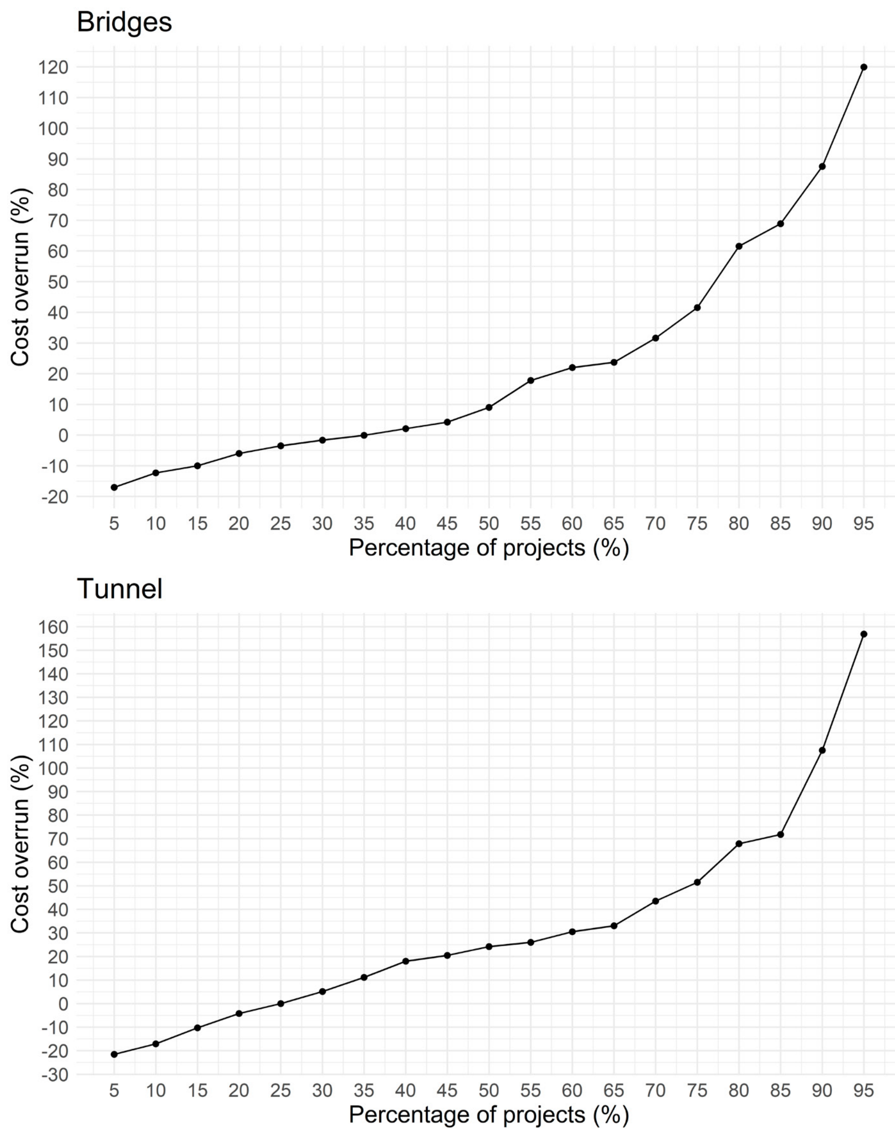


Figure 9: RCF of cost overruns in international bridge and tunnel projects

6.2 Schedule Overrun

The RCF curves in Figure 10 and Figure 11 show the cumulative probability distribution of schedule overrun values within the P5 – P95 percentile interval in international data. 50% of the roadway projects had a schedule overrun of 11% or less, and 80% of the roadway projects had a cost overrun of 60% or less. Furthermore, 50% of the railway projects had a schedule overrun of 17% or less and 80% of the railway projects had a cost overrun of 52% or less.

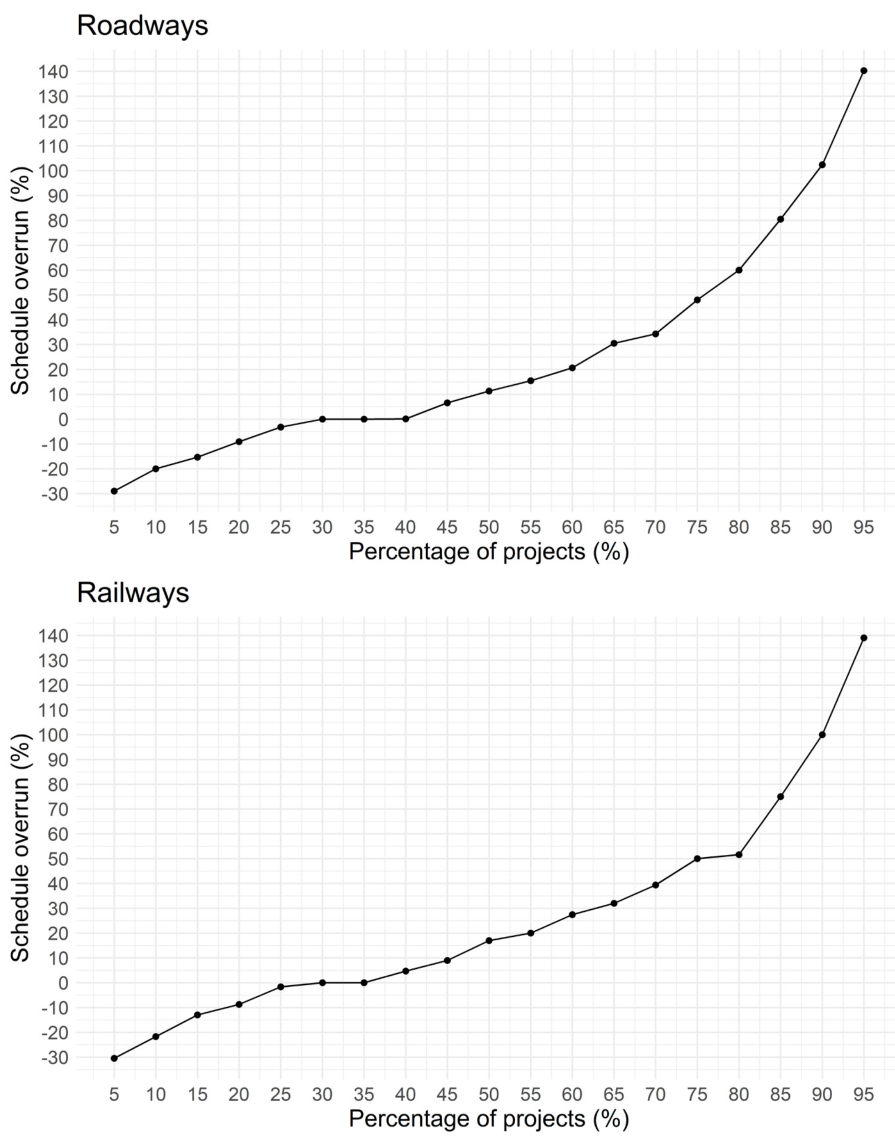


Figure 10: RCF of schedule overruns in international roadway and railway projects

50% of the bridge projects had a schedule overrun of 16% or less, and 80% of the bridge projects had a cost overrun of 39% or less. Furthermore, 50% of the tunnel projects had a schedule overrun of 1% or less and 80% of the tunnel projects had a cost overrun of 42% or less.

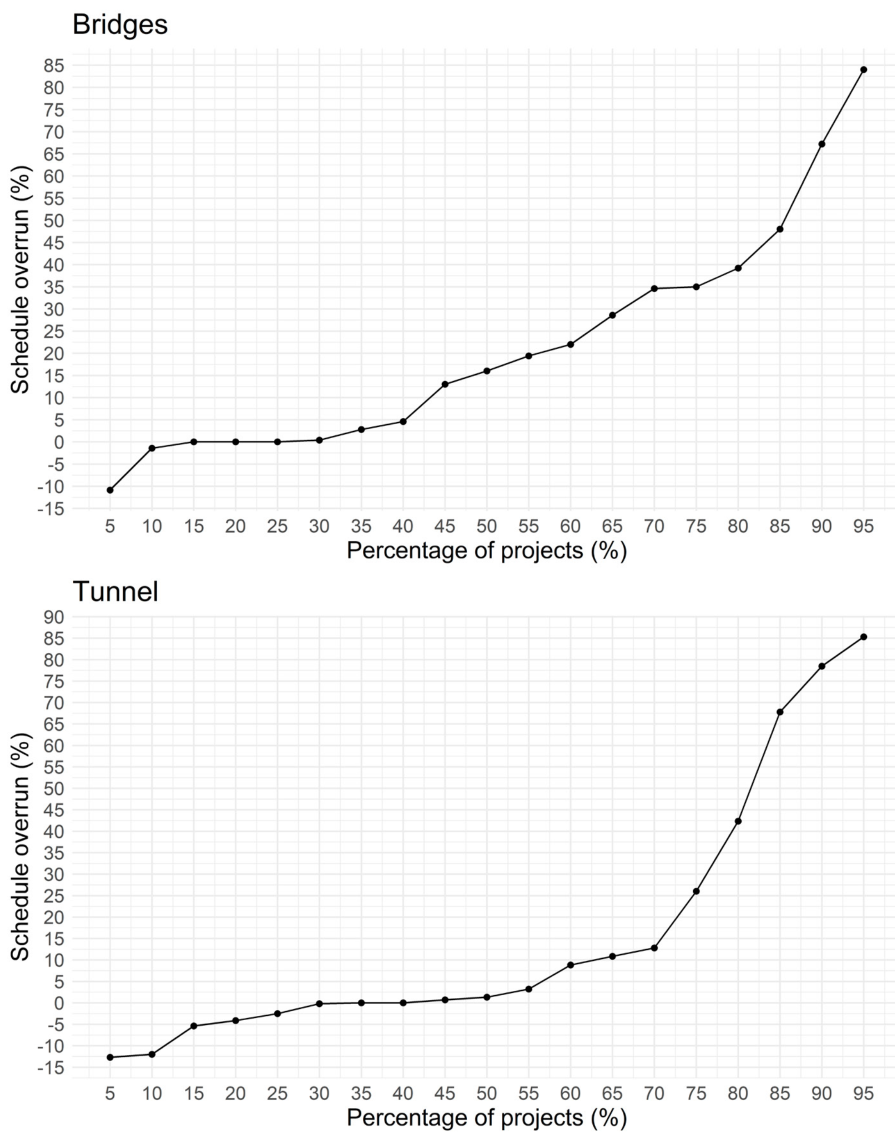


Figure 11: RCF of schedule overruns in international bridge and tunnel projects

6.3 Unit cost

The RCF curves in Figure 12 and Figure 13 show the cumulative probability distribution of cost per lane kilometer values within the P5 – P95 percentile interval in international data. 50% of the roadway projects had a unit cost of 6.4 million BRL or less, and 80% of the roadway projects had a unit cost of 12.1 million BRL. Furthermore, 50% of the railway projects had a unit cost of 42.7 million BRL or less and 80% of the railway projects had a unit cost of 107.7 million BRL or less.

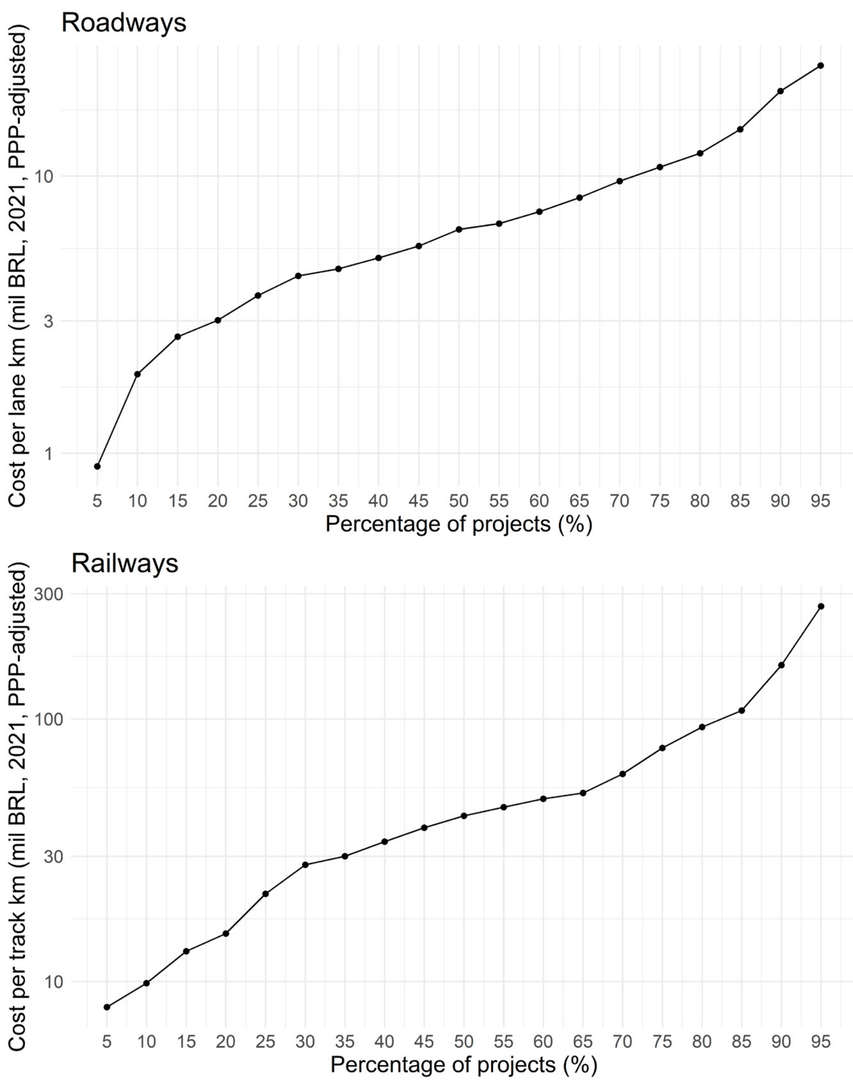


Figure 12: RCF of cost per kilometer in international roadway and railway projects

50% of the bridge projects had a unit cost of 103.8 million BRL or less, and 80% of the bridge projects had a unit cost of 203.7 million BRL. Furthermore, 50% of the tunnel projects had a unit cost of 46.5 million BRL or less and 80% of the tunnel projects had a unit cost of 147.3 million BRL or less.

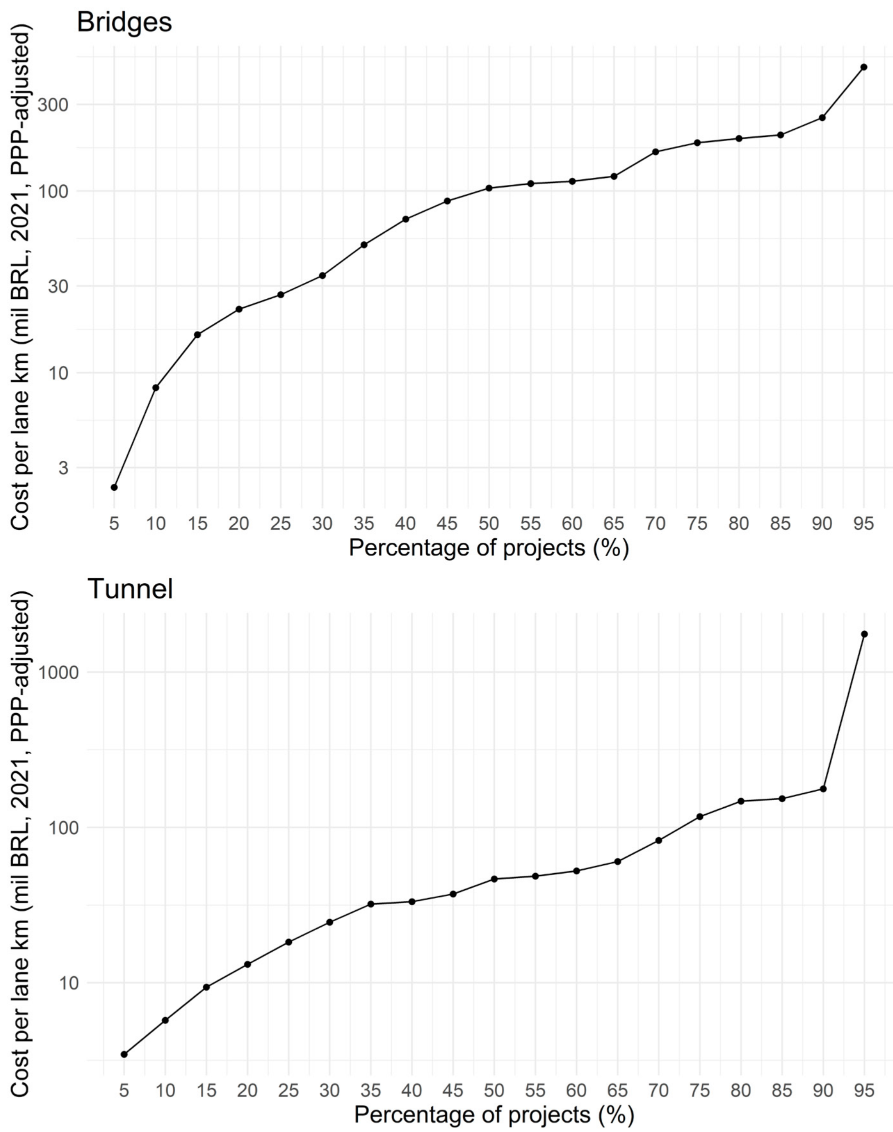


Figure 13: RCF of cost per kilometer in international bridge and tunnel projects

7. GUIDE ON HOW TO USE RCFs

The project management profession is notoriously poor at forecasting how projects will perform on budget, time, and benefits. In the world's largest academic quality dataset on projects (in excess of 17,000), less than half of them are completed within budget, 8.5% are completed within budget and schedule and just 0.5% are completed within budget and schedule and deliver the expected benefits or more[9]. Professor Bent Flyvbjerg coined the expression The Iron Law of Projects: 'projects are over budget, over time, under benefits, over and over again'[10]. This is seen on a wide range of project types, from the Olympics to IT and transport.

	Solar power	Roads	Rail	Buildings	IT-led change	Dams	Olympics	Nuclear waste storage
Cost overrun	1%	16%	39%	62%	73%	75%	157%	238%
Frequency of cost overrun	4 of 10	6 of 10	7 of 10	7 of 10	4 of 10	7 of 10	10 of 10	9 of 10
Schedule overrun	2%	36%	32%	32%	43%	44%	0%	70%
Benefits overrun	n/a	-5%	-23%	-5%	-28%	-11%	n/a	-23%
Cost Black Swans	0%	4%	10%	20%	18%	23%	57%	43%
Ø duration, years	2.2	4.1	8.0	7.9	3.3	8.0	7.1	6.8

Figure 14: Project Performance Map[11]

7.1 Theory

Historically, explanations for inaccurate project forecasts have focused on limited data quality or imprecise modelling. However, if this were the case errors would cancel themselves out across projects showing, for actual performance, a normal distribution of outcomes where the number of times that forecasts were optimistic being roughly equivalent to the times that they were pessimistic. However, analysis of actual outcomes presents a very different picture with a distinct skew to negative outcomes as illustrated below (and encapsulated by the Iron Law of Projects).

9 Source: Oxford Global Projects Database (Q4 2022).

10 Flyvbjerg, B, *From Nobel Prize to Project Management: Getting Risks Right.* Project Management Journal, 2006 vol. 37, no. 3, pp. 5-15

11 Ibid. Note: Measured from date of decision to build, in constant prices. Black Swans are defined as statistical outliers.

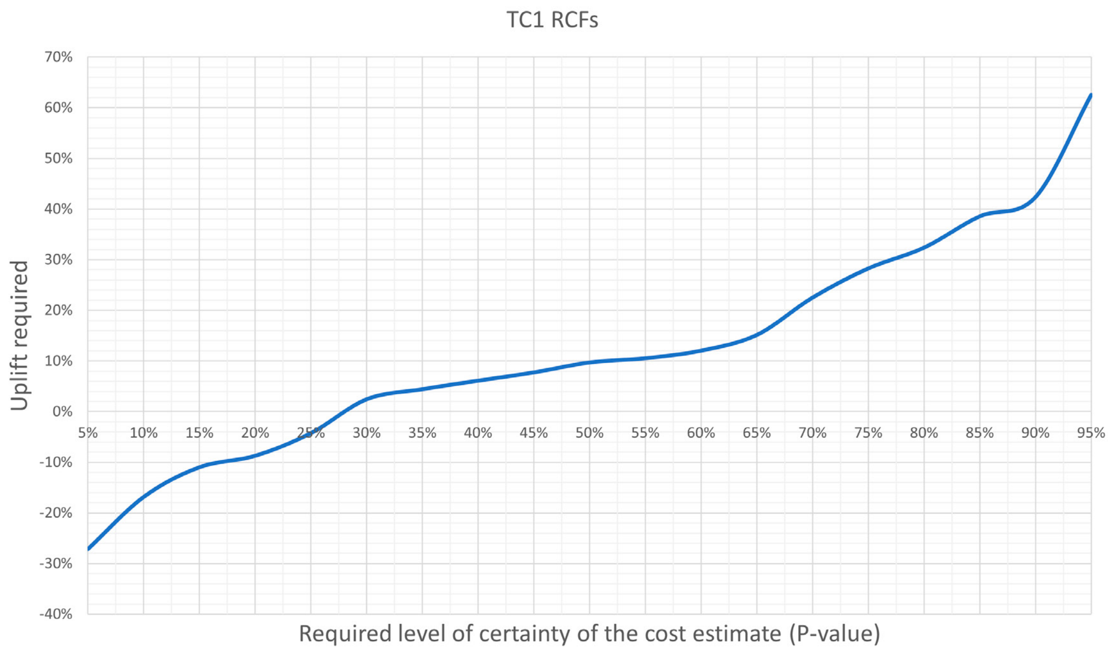


Figure 15: Reference Class Curve for actual performance of roads in Ireland [12]

This sample diagram above from actual road data across Ireland illustrates a best case outcome of -25% but a worst case of +60%. An asymmetric skew with the worst case more than twice the best case.

A far more robust explanation for the poor project forecasting we see lies in the Nobel-Prize winning work of Amos Tversky and Daniel Kahneman. They posited:

> *"the **planning fallacy**; a consequence of the tendency to neglect distributional data, and to adopt what may be termed an 'internal approach' to prediction, where one focuses on the constituents of the specific problem rather than on the distribution of outcomes in similar cases."*[13]

This is very pertinent for project forecasting, as estimates (cost, time and potential benefit) often summate individual elements of a project commonly referred to as an 'internal' or 'bottom up' approach. This approach has limited ability to acknowledge strategic factors, the interaction of individual elements (eg complexity) or 'unknown unknowns' which can be seen to have affected similar projects[14]. Whilst it may consider historic costs at the elemental level the total project performance is not considered.

12 RCF Guidelines for National Road Projects, Transport Infrastructure Ireland.

13 Kahneman, D., Tversky, A., "Intuitive prediction: Biases and corrective procedures," *TIMS Studies in Management Science*, 1979, Vol 12, pp. 313–27.

14 Bartlett, M and Leed, A, Independent Risk Evaluation, *European Transport Conference* 2022

Internal approaches are highly prone to biases[15] with even experts, who are aware of potential biases still subject to them[16]. One of the most prevalent is optimism bias recognised in, for example, the United Kingdom's Treasury guidance since 2004[17]. This is a systematic tendency to assume that outcomes will be better than they are likely to be and is manifested as underestimating the likelihood of negative events and overestimating the likelihood of positive occurrences. It is easy to see how this would result in worse project outcomes than anticipated being more likely than improved outcomes.

In addition to optimism bias affecting the robust ranging of estimates, political bias also offers an explanation for 'strategic misrepresentation':

> *"**Strategic misrepresentation** is the planned, systematic distortion or misstatement of fact – lying - in response to incentives in the budget process."*[18]

Both national or multi-national 'Political' and internal or smaller scale 'political' considerations can incentivize a misrepresentation of reporting and forecasting. It is not unknown for senior leaders or stakeholders to announce a completion date and budget for a high-profile project before those involved in delivering the project have a fully defined scope or a reasonable knowledge of what will be involved in delivering it.

Whilst optimism bias and political bias both play a role in explaining poor project forecasting, it is difficult to estimate the actual impact of either or both by the 'internal' approach. It might be expected that the more high-profile a project is or the more a project is the 'pet project' of a senior leader, the more likely that political bias will be significant.

RCF is comprised of three steps:

1. Identify a relevant reference class
2. Establish a probability distribution for the reference class
3. Make the forecast

1. Identify a relevant reference class

This involves finding similar projects that have been completed and determining how they performed versus expectations at key decision points. Relevant project parameters are gathered to reasonably correlate with historic data. At least 15-20 projects should be used.

15 Kahneman, D., Thinking Fast and Slow, Farrar, Straus & Giroux, 2011

16 Tetlock, Philip and Gardner, Dan. Superforecasting: The Art and Science of Prediction. Penguin Random House, 2019.

17 HM Treasury, The Green Book: appraisal and evaluation in central government, 2004

18 Jones, L.R. ; Euske, K.J., "Strategic Misrepresentation in Budgeting", Oxford University Press, *Journal of Public Administration Research and Theory*, 1991, Vol.1 (4), p.437-460.

This report has produced a series of reference classes for cost overrun (cost risk), schedule overrun (schedule risk) and unit cost (cost per km) based on historical projects which have been completed, and for which credible cost and schedule data was available for both estimates and outturn. These reference classes have been constructed for the following project types:

- Roads
- Rail
- Tunnels
- Bridges

If used for forecasting, simply choose the reference class that fits your project in terms of project type and type of forecast.

2. Establish a probability distribution for the reference class

This stage determines how the variable in question performed in comparison to the estimate for each project in the Reference Class.

A cumulative distribution is then created by sorting the performance data from the largest to smallest overrun and then the relative share of each data point in the sample is calculated (e.g. if 25 projects are in a reference class each project has 4% share) and summed up so that the distribution ranges from 0%-100% (i.e., the project with the largest overrun project represents 4% the second highest overrun 8% and so on as illustrated in the following figure).

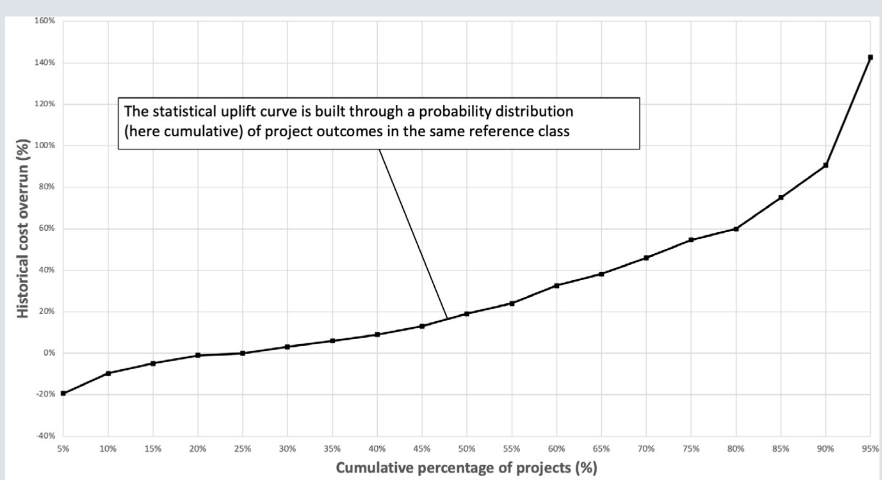


Figure 16: Cumulative probability distribution of overrun in the reference class (conceptual)

The reference classes developed for this report have all been displayed as cumulative distributions - or “Reference Class Curves”.

3. Make the forecast

The final step is to review the cumulative distributions and identify the necessary uplifts to de-bias the estimates. For this the curve is reinterpreted. The cumulative percentage of projects with a given overrun in the reference class now becomes the acceptable chance of overrun and the uplift to add to the base estimate to de-bias it; a consideration of *risk appetite*.

For example, if decision makers accept a 50% chance of overrun (i.e. they require a 50% certain estimate or P50) then the relevant uplift at the 50% position is added. If decision makers are more risk averse and only accept a 20% chance of overrun (i.e. they require a 80% certain estimate or P80) then the uplift at the 80% position is utilised.

P50 is a common position taken by organisations using risk analyses and/or RCF to set a project *target cost*.

P70 to P90 is the range used by these organisations to determine a more pessimistic forecast. However, it is not always deemed pragmatic to actually set aside contingency or float (schedule) to this level as other decisions (such as de-scoping) may be feasible.

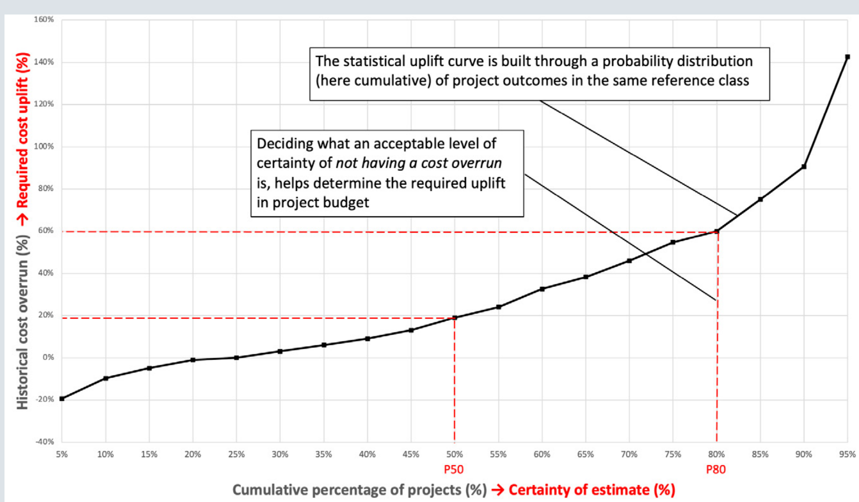


Figure 17: Establishing the uplifts as a function of the acceptable chance of cost overrun based on the cumulative distribution of cost overrun in the reference class (conceptual)

7.2 Integrating RCF with Bottom-Up Risk Estimation Techniques

Reference Class Forecasting (RCF) can complement bottom-up risk estimation approaches, such as Quantitative Risk Analysis (QRA), to provide a more robust and comprehensive understanding of potential project outcomes. QRA is a probabilistic technique used to assess the potential impact of identified risks and uncertainties on a project's objectives, by quantifying their likelihood and consequences. While QRA typically generates a narrower range of estimates, focusing on specific project risks and uncertainties, RCF offers a broader, more pragmatic perspective based on historical data from similar projects. By comparing the two methods, project managers can identify potential gaps or biases in their bottom-up risk assessments and develop a more accurate forecast of project performance. Furthermore, the combination of RCF and QRA can help establish a holistic risk management strategy that accounts for both project-specific risks and the wider context of similar infrastructure projects.

In the UK Infrastructure and Projects Authority's Risk Management Module of the Project Routemap, the integration of RCF and QRA is recommended as a valuable approach to improve project forecasts and risk management. According to the module, "the combination of bottom-up risk management processes and top-down, evidence-based forecasting will help to identify project-specific risks and uncertainties that may not have been considered, and will provide a more reliable basis for decision-making"[19]

It should be noted that the RCF distribution is based on the historical overruns in similar, completed projects. Thus, projects might need to consider whether any additional adjustments to the chosen level of certainty (P-level) are needed or whether the RCF should only apply to parts of the base estimate. In other words, whether the project at hand is more or less risky than past projects. Examples for deviation could be:

- If a project has progressed further with a detailed design development at a given stage than projects normally would have.
- If all necessary land had already been acquired, then there would be no need to apply an uplift to this element of the work
- If the financial risk has been fully transferred to a subcontractor; however, this would need to be exercised with caution to make sure the risk is fully financially transferred through a mechanism such as a fixed firm price with no potential for claw backs. Keep in mind that a project will never be risk free if it has not yet completed even with an extensive contract in place. Thus, even at this stage, consideration should be made of the potential for strategic non-transferrable risks and design changes which could still affect the outturn.

Any adjustment to the application of RCF ought to be based on <u>hard evidence</u> in order to avoid reintroducing optimism back into the estimate.

19 UK Infrastructure and Projects Authority. (2022). Project Routemap: Risk Management Module. Retrieved from ‘https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1080243/Risk_Management_-_FINAL.pdf’

8. RECOMMENDATIONS

8.1 Data Quality and Transparency

The report discusses the challenges in drawing meaningful insights from comparisons between Brazilian datasets and international project level data due to deficiencies in the entire workflow of data capture, management, and processing. These deficiencies encompass several aspects that warrant further development and improvement, including identifying the necessary data to be collected, implementing effective data collection, processing, and consolidation in a centralized manner, using better storage methods (preferably through electronic systems to promote digitalization and accessibility), promoting transparency and accessibility, and providing instructions for interpretation to increase usability. Transparency is a significant issue within this broader context, as the lack of transparent data collection and reporting frameworks creates obstacles in drawing meaningful conclusions from the currently collected data. These obstacles arise due to difficulties in matching work contracts to their respective projects, tracking cost and schedule overruns, and identifying relationships between new contracts and previous contracts.

Addressing these data management deficiencies is crucial not only for Brazil but also for other countries in Latin America and the Caribbean that may face similar challenges. By sharing best practices and lessons learned in improving data management processes, countries in the region can collectively enhance the quality and usefulness of infrastructure project data for decision-making and performance evaluation. This can lead to more efficient and effective management of infrastructure projects, better allocation of resources, and reduced risks of corruption and unfinished public works. Furthermore, by improving data management processes throughout the region, countries will be better equipped to share and compare data across borders, leading to more informed decisions and better overall outcomes for infrastructure projects.

The current Brazilian datasets further lack transparency in terms of identifying the amount of work completed by each contract of a project, making it difficult to calculate unit costs accurately. The variable "Extension (km)" does not provide clear information on the actual number of kilometers of road or rail that have been laid in each contract. Therefore, it is recommended that Brazilian agencies (DNIT and Valec) improve their data collection methods to capture more detailed information on the actual work completed by each contract, including the physical characteristics of the infrastructure such as the number of lanes/tracks, width of the road, etc. By capturing this information, it will be easier to calculate unit costs accurately and find a proper reference class for comparison purposes.

To address the limitations in Brazilian data, it is recommended that data capture practices be improved to encompass the entire project, not just the construction share. This applies not only to data collection practices in Brazil but to other OLACEFS as well. These guidelines can be used together with international benchmarks to improve OLACEFS service offerings for the active participation of AIs in national anti-corruption systems, increase the scope of AI audits, and strengthen the exchange of technical cooperation and collaboration of CSOs and other government agencies.

Currently, the lack of control and transparency in Brazilian data inhibits efforts to improve Brazilian infrastructure project management and reduce political corruption in Brazil. To overcome this challenge, OGP recommends transparent and comprehensive data capture: Data collection should be done at both the project and contract levels, with contracts clearly attributed to their respective projects. This will enable more accurate calculations of cost and schedule overruns and improve transparency in the allocation of project costs across contracts. Data capture should encompass the entire lifespan of the project, not just the construction phase, allowing for more comprehensive assessments of project performance and better monitoring of project costs. Additionally, subsequent contracts should reflect the original cost and schedule estimates of the project, as well as any deviations from these estimates, to prevent artificial underestimations of cost and schedule overruns.

Finally, to promote greater accountability and transparency in the procurement and execution of government contracts, the report recommends that the Brazilian government should consider implementing a mechanism to ensure accurate reporting, tracking, and monitoring of cost and schedule overruns. This could include clear and transparent documentation of the original and subsequent contracts, changes to the scope or timeline of the project, and regular reporting on progress and cost overruns.

When a project is divided into many contracts during execution, it can be challenging to keep track of costs. To manage this, governmental agencies often require contractors to provide regular progress reports, including information on costs incurred to date. This information is then used to update the overall project cost data. In addition, many agencies require detailed cost breakdowns for each contract, which can help to ensure that all costs are accounted for and properly allocated. Project management software and cost accounting systems can also be used to track costs across multiple contracts and ensure that all costs are accounted for in the overall project budget.

The report also suggests that penalties or consequences could be imposed on contractors who fail to accurately report cost and schedule overruns and on government officials who fail to enforce reporting requirements. By implementing these measures, the government can improve the efficiency and effectiveness of infrastructure projects in Brazil and reduce the risk of political corruption and unfinished public works.

By implementing these recommendations, OLACEFS can strengthen its efforts to combat corruption and improve infrastructure project management in Brazil and other SAI member countries.

8.2 Employing Reference Class Forecasting

In addition to improving data quality and transparency, it is essential to utilize effective estimation methods and risk management strategies for infrastructure projects. The RCF curves developed in this report offer a powerful approach to predicting project outcomes and mitigating risks by drawing upon historical data and experiences from similar projects. This section outlines additional recommendations on how to effectively use the RCF curves. By applying these recommendations, project managers and decision-makers can enhance their ability to accurately forecast project outcomes and implement risk management strategies that drive project success, reducing optimism bias and strategic misrepresentation.

1. **Balancing Top-down RCF and Bottom-up QRA Approaches**: As project development progresses, the appropriate method for calculating risk exposure should change. In the early stages of a project, when there are many uncertainties and potential routes for development, a top-down RCF approach is more suitable. As the project nears completion, a bottom-up QRA approach becomes more applicable due to the availability of detailed information and fewer remaining uncertainties. The degree of project definition and organizational maturity will determine when the transition from top-down RCF to bottom-up QRA occurs.

2. **Ensuring Adequate Project Design and Detailed Bottom-up Budgeting**: While RCF can be a valuable tool for managing risk exposure during project development, it should not replace the need for detailed bottom-up budgeting and thorough project design. A common issue contributing to paralyzed works, cost overrun, and schedule overrun of infrastructure projects in Brazil is the insufficiency or inadequacy of feasibility studies and project designs. It is crucial for project owners and decision-makers to invest sufficient effort in designing and detailing project costs to minimize risks and uncertainties effectively. It is important to recognize that relying solely on RCF may lead to a false sense of security regarding risk coverage, which could, in turn, result in suboptimal project design and inadequate bottom-up budgeting.

3. **Using RCF Uplifts for Robust Business Cases and Portfolio-level Contingency**: RCF uplifts should not be added directly to a project's budget as this might lead to overspending. Instead, they should be used to confirm that a business case remains robust if costs rise to the uplifted level and to create a contingency held at the portfolio level. This approach ensures better financial management and risk mitigation.

4. **Implementing "Skin in the Game" Incentive Structures**: To prevent perverse incentives that could arise from expecting contractors to exceed budgets and deadlines, it is crucial to establish incentive structures that align contractor interests with project targets. Contracts should be structured in a manner that rewards contractors for meeting targets and penalizes them for falling short. This approach, known as "skin in the game," ensures that contractors are motivated to deliver projects on time and within budget, while also providing a safeguard against potential misuse of RCF-based expectations.

5. **Continuously Update RCF Data with Regional Inputs**: As Brazil and other countries in the region improve their data collection processes, it is crucial to periodically update the RCF model with newly available data. Incorporating accurate and reliable data from Brazil and neighboring countries will enhance the model's predictive accuracy and adaptability to local conditions. This ongoing process of refining the RCF model ensures that forecasts remain relevant and reliable, ultimately contributing to better-informed decision-making and more effective risk management strategies in infrastructure projects. By fostering a collaborative environment among regional stakeholders and prioritizing data-driven approaches, infrastructure projects can benefit from enhanced performance and reduced risks of cost and schedule overruns.