Assessment of unit costs (standard prices) of rail projects (CAPital EXpenditure)

Final report

Contract No 2017CE16BAT002

Written by M. ATTINÀ, A. BASILICO, M.BOTTA, I. BRANCATELLO, F. GARGANI, V. GORI, F. Wilhelm, M. MENTING, R. ODOARDI, A. PIPERNO, M. RANIERI from PwC and supported by the contribution of Prof. M. COLI and Prof. F. ZAMBELLONI from COWI

13th July 2018

EUROPEAN COMMISSION

Directorate-General for Regional and Urban Policy
Directorate F - Closure, Major Projects and Programmes Implementation III
Unit F1 - Competence Centre Closure and Major Projects

Contact: Jonathan DENNESS

E-mail: REGIO-F1-HOU@ec.europa.eu

European Commission
B-1049 Brussels

Assessment of unit costs (standard prices) of rail projects (CAPital EXpenditure)

Final report

Contract No 2017CE16BAT002

Directorate-General for Regional and Urban Policy

2018 EN

LEGAL NOTICE

Luxembourg: Publications Office of the European Union, 2018

ISBN: 978-92-79-91813-1
doi: 10.2776/296711

The study has been prepared by (in alphabetical order):

Massimo ATTINÀ (PwC), Andrea BASILICO (PwC), Matteo BOTTA (PwC), Irene BRANCATELLO (PwC), Francesco GARGANI (PwC), Valerio GORI (PwC), Frédéric Wilhelm Maria MENTING (PwC), Roberta ODOARDI (PwC), Alessandro PIPERNO (PwC), Marco RANIERI (PwC) and was supported by the contribution provided by various experts, including Prof. Massimo COLI (COWI) and Prof. Franco ZAMBELLONI (COWI).

This document has been prepared for the European Commission DG-REGIO (the Client) in accordance with the Terms of Reference 2016 CE 16 B AT 087 under the Multiple Framework Service Contract 2013CE16BAT0642012 Lot 1 and solely for the purpose and on the terms agreed with the EC. We accept no liability (including for negligence) to anyone else in connection with this document or any other outcome of the assignment.

This document and the REGIO Rail Unit Cost Tool contain information obtained or derived from a variety of third party sources as indicated within the document. PwC has not sought to establish the reliability of those sources or verified the information so provided.

Should any person other than the European Commission DG-REGIO obtain access to and read this document or have access and use the REGIO Rail Unit Cost Tool, such persons accepts and agrees to the following terms:

1. The reader of this document as well as the user of the REGIO Rail Unit Cost Tool understands that the work performed by PwC was performed in accordance with instructions provided by our client, and was performed exclusively for their benefit and use. The document may therefore not include all matters relevant to the reader and the REGIO Rail Unit Cost Tool may not include all functions relevant to the user. The document and the REGIO Rail Unit Cost Tool are provided for general guidance only and do not constitute investment or any other advice.

2. The reader agrees that PwC accepts no liability (including for negligence) to them in connection with this document and the related REGIO Rail Unit Cost Tool.

PwC was commissioned to develop a tool to support the evaluation of the investment cost of rail infrastructure projects, taking into account the factors impacting on the cost, and provide indication on the likely final cost of the projects. The evaluation performed by the tool consists of a statistical analysis of the unit cost of an existing investment with the unit cost of comparable projects included in the database connected to the tool. Data on these projects have been retrieved from desk research, literature review and provided by the stakeholder involved in the study. The results of the evaluation do not represent a measure of the quality of the investment assessed or of the cost estimates. They only present the positioning of the unit cost of the investment assessed within the distribution of unit cost of similar projects.

Table of Contents

List of figures

List of tables

Glossary

Abbreviations	Definition
CEF	Connecting Europe Facility
DG REGIO	Directorate-General for Regional and Urban Policy
DTU	Danmarks Tekniske Universitet
EC	European Commission
ERDF	European Regional Development Fund
ERTMS	European Rail Traffic Management System
Estimated cost	Rail investment cost reported in the feasibility study or assessed in the planning & design phase
ETCS	European Train Control System
EU	European Union
Final cost	Cost reported in the contracts concerning the implementation of the rail project
IM	Infrastructure Manager
INEA	Innovation and Networks Executive Agency
MA	Managing Authority
MoT	Ministry of Transport
MS	Member State
OECD	The Organisation for Economic Co-operation and Development
OLS	Ordinary Least Squares
PPP	Purchasing Power Parity
TEN-T	Trans-European Network - Transport
Tier 1	First breakdown of project characteristics and costs. Project are categorised based on project general factors (e.g. line category, areas of the project, geographical scope) and costs are split into three sub-categories, namely planning & design, construction works and ancillary costs
Tier 2	Second breakdown of project characteristics and costs. Technical information relative to the rail infrastructures envisaged in the project (e.g. tunnels, bridges, and viaducts) is included. Costs for construction works and ancillary are split into further subcategories.
Tier 3	Third breakdown of project characteristics and costs, focusing on costs and technical characteristics relative to the equipment and permanent way, and civil engineering works.
TSI	Technical Specifications for Interoperability
UIC	*Union Internationale des Chemins de Fer* - International Union of Railways
WP	Work Package

Executive summary

Background and scope of the study

Bringing the people and economies of the individual Member States closer by developing a functioning, efficient and interconnected European transport system is a paramount objective of the European Union. [1] In particular, the European rail network poses huge opportunities, as rail is an environmentally friendly and efficient transport mode. Therefore, the European Commission is investing in the development of the rail network to increase the share of rail in freight and passenger transport. Nonetheless, the demand for funding exceeds the public resources available, thus it is imperative to improve the efficiency of public procurement and planning of public funds, in order to improve the efficiency in terms of impact of the available funds.

Several studies, carried out in the past, highlighted the difficulty of assessing the delivery efficiency of capital investments in rail infrastructures due to the various interacting elements involved, such as the project features, the technical complexities and the political, regulatory and natural differences that can be found across the Member States.

Aware of such difficulties, the Authors undertook a wide research about the relationship between project costs and technical features, in order to build a reliable and user-friendly tool. The main aim of such tool is to aid the various decision makers involved the infrastructure planning process, as the European Commission, Managing Authorities, Contracting Authorities, JASPERS, Infrastructure Managers, Strategic Rail Authorities, etc. in the assessment of the investment cost of rail infrastructure projects. Following extensive work, the authors developed the REGIO Rail Unit Cost Tool, based on a review of the project unit costs in the rail sector in core and comprehensive sections of the Trans-European Transport Network, and assessing factors that lead to the variation of the costs of infrastructure development. The tool provides a timely benchmark of each rail project costs compared to a large sample of projects with similar characteristics. At the same time, using the same logic, the tool can provide indications of the average cost of infrastructure with certain characteristics, so to provide support for the forecasting and planning of rail investments.

In this exercise, the Authors have been supported by an extensive network of stakeholders. These have been particularly relevant in collecting the data on projects as well as in supporting the definition of the features the REGIO Rail Unit Cost Tool had to incorporate. A dedicated Steering Group (SGUCR) of experts actively contributed to define the scope of the assignment and monitoring and directing its progress.

The SGUCR provided important information on different practices across MS and approved the key deliverables at each stage of the assignment. Its contribution allowed to fine tune the methodology and apply the necessary correction to the text presented, including support on the definition of the applicable technical standards among the different Member States.

---

[1] Among others:

European Commission 2016 Communication to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, *A European Strategy for Low-Emission Mobility* COM(2016) 501 final

European Commission 2016 Staff Working Document, *The implementation of the 2011 White Paper on Transport "Roadmap to a Single European Transport Area – towards a competitive and resource-efficient transport system" five years after its publication: achievements and challenges* SWD(2016) 226 final

European Commission 2011 WHITE PAPER Roadmap to a Single European Transport Area – Towards a competitive and resource efficient transport system COM(2011)0144 final

The REGIO Rail Unit Cost Tool represents the practical tool developed to respond to the specific objectives of the study, which are:

- to establish common understanding of factors influencing rail infrastructure investment cost,
- to elaborate a methodological approach to identify and calculate unit cost in rail infrastructure investments,
- to provide input to decision-makers for high-level evaluation of the costs of rail infrastructure programmes and projects,
- to assist authorities in charge of rail infrastructure programmes and projects to communicate effectively to a wider set of stakeholders about project cost.

The results of the study will enable public stakeholders involved in the procurement process to increase transparency and awareness on unit costs, hence to early identify possible issues of overpricing or under-pricing of rail infrastructure projects.

Main differences with previous studies

Several initiatives have been undertaken over time with the aim of enhancing the planning, design and procurement phases of railway projects by identifying unit cost of rail infrastructure projects. The authors undertook a process of analysis and scrutiny of past studies and projects to synthesise all the different approaches and learn from past experiences. The analysis was focused on specific studies on rail unit costs with a global scope and with a cost data time range of over 25 years. The literature review allowed to gain a solid and clear understanding of the factors which resulted in a significant variability in the unit cost identified.

The analysis of previous studies highlighted how the limited amount of data available was one of the causes that limited the accuracy of the analysis of the unit cost. Therefore, a strong team composed of several rail experts and dedicated team operating on-site have been involved in the analysis, collecting information from the Infrastructure Managers over Europe, for a total effort of over 450 person/days.

In order to achieve a statistically significant population of data on rail projects, which goes beyond the amount of data previous studies had access to, this present study combined information that was available only from DG REGIO and INEA databases in form of project funding applications with information directly retrieved from infrastructure managers by PwC experts in the local language and anonymised propriety information from PwC databases on prior projects and studies. Through the combination of these resources with other information available in specialist publications or in general in the public domain, the authors of the study had access to more information than previous studies were able to cover.

The other factor impacting the accuracy of the unit cost identified in previous studies was the limited uniformity of the rail projects analysed in terms of investment characteristics. The authors of the study performed detailed analysis on the investment features which impact on cost of rail infrastructure to create a well-defined sample for the analysis. Further, a breakdown of the investment was elaborated to collect detailed information on the different components of the each investment. This led to the creation of the structure of the project database, where information on technical features and cost of over 200 rail infrastructure projects was finally stored.

In the literature, different ways to investigate infrastructure unit costs was attempted. These fall into two dichotomous approaches. The first aims to identify a statistical, universally applicable law linking the construction costs and the features and factors characterising the infrastructure. The effectiveness of a similar approach is dependent on the possibility to assess a very large amount of observations through regression analyses. The other approach aims to investigate the relationship between cost and characteristics of a rail infrastructure by clustering comparable projects and costs and analysing them separately. Compared to the previous approach, the latter requires less observations to produce significant and reliable results – more pertinently, to produce the dataset known as the critical mass of statistical data. Considering the number of data available at disposal, the authors adopted this latter approach.

THE DATABASE

The population of the database entailed a significant effort in terms of both data collection and cleaning. The database was initially built based on the data made available by DG REGIO and INEA on over 860 projects. This was accurately analysed to retrieve information on technical features and cost of the projects with the aim of creating a detailed base for the analysis and exclude out-of-scope investments (i.e. not related to works performed on a railway line).
The process required the analysis of approximately ten to 30 documents per project to be investigated with up to 300 pages for individual documents. This activity involved ten experts who were assigned to this task over a three months period.

Besides this extensive desk analysis, the authors of the study undertook considerable research to complement the information collected with data retrieved from different studies, analysis, essays available on the internet. Furthermore, IMs from all the EU Member States have been contacted to gather detailed information on their investments. Considering the effort required, on-site teams have been activated to support them.

The whole process enabled the creation of a database, containing detailed information on over 200 rail infrastructure projects developed in the EU over the 2000 – 2015 period. The database itself is a relevant outcome of the study, given its size – which is significantly larger than those of the past studies on the matter – and the level of accuracy – for most projects, information has been collected on each single work component included in the investment.

Such project mapping was a prerequisite to the development of the REGIO Rail Unit Cost Tool, the digital tool designed to support the assessment of investment cost of rail infrastructure project based on a comparison with the investment cost of similar projects.

THE REGIO RAIL UNIT COST TOOL

The main objective of the REGIO Rail Unit Cost Tool is to perform a detailed benchmarking analysis in order to compare the unit cost of rail infrastructure investments with those of projects presenting similar characteristics, enabling to identify investment costs which are in line with those of past projects or deviate from the usual values.

To identify the projects examined in the benchmark analysis, the REGIO Rail Unit Cost Tool considers all the cost drivers identified during the data definition process, namely types of works included in the project, infrastructure category, design speed, topography and category of service of the line among others such as dedicated passenger/mixed use/ freight and selects the investments which, as for these drivers, present similar characteristics to the project investment under analysis.

To ensure the cross-country validity of the analysis, all the investment costs are normalised following a specific index, developed based on the PPP and the Producer price index. The normalisation takes into consideration the differences between the economies of the EU MS and the differences over time. Yet, it cannot embraces of the different aspects of the rail infrastructure market of the EU countries. These were analysed and reported in country sheets that complement the outcomes of the study and the evaluation performed by the REGIO Rail Unit Cost Tool.

The main distinctive feature of the REGIO Rail Unit Cost Tool is its "self-learning" ability. Indeed, once the analysis of a specific project is completed, the user has the possibility to include in the database the data of the project assessed. Therefore, the cost ranges considered in the analysis performed by the REGIO Rail Unit Cost Tool update as the tool is used. Such a feature ensures that the accuracy and the quality of the analysis performed will constantly improve in the future, provided that investment costs are correctly input. DG REGIO ensures that resources will be made available to feed the database of the REGIO Rail Unit Cost Tool with additional data and thereby update the tool over time. To ensure correctness and comprehensiveness of data input, is likely to require the dedication of some resources will have to be dedicated by DG REGIO.

The maintenance and update of the REGIO Rail Unit Cost Tool is necessary to ensure its correct functioning in the following years. It is vital to ensure that the projects included in the database are relevant and correctly input as well as that macroeconomic parameters to compare costs across time and countries are periodically updated. The REGIO Rail Unit Cost Tool is dependent on the information that are stored in its database. Would the database be populated with odd observations, the results would be equally odd.

The REGIO Rail Unit Cost Tool is designed to support evaluators of projects to have indications of the positioning of single projects against the usual unit cost values of similar projects (at overall and component level). The tool can therefore be used in any project assessment exercise, i.e. during the project option analyses or analysis of costs of project proposals.

The tool is further designed to extrapolate statistical distributions from the observations in the database to support the forecasting of infrastructure investments. Users can input the main characteristics of the desired infrastructure to study and instruct the tool to return the average costs that projects with similar characteristics had. The tool is in no case designed to substitute the expertise and the judgement that the evaluators provide.

The tool presents several opportunities for further development A number of potential development of features to enhance and widen the application of the REGIO Rail Unit Cost Tool is provided in Section 6.

QUALITATIVE AND QUANTITATIVE RESULTS OF THE STUDY

The results of the analysis performed by the REGIO Rail Unit Cost Tool enabled to validate the approach followed in the construction of the tool and represent a relevant output of the study per se.

The information included in the database has been analysed following the approach implemented in the REGIO Rail Unit Cost Tool. Unit cost ranges have been determined by clustering projects and costs based on:

- the different components of a rail infrastructure projects. i.e. the investment cost of rail infrastructure project is breakdown into its subcomponents and the unit cost range is identified for each of them;
- the internal and external factors that impact on cost of rail infrastructure projects (e.g. orography, urbanisation, market conditions, etc.).

Following this approach, it was possible to identify the total investment unit cost ranges for the different rail infrastructure categories and type of intervention on the infrastructure with an increasing level of detail and referred to in more narrow clusters.

The organisation of the information was designed through a three-tier approach. Each tier corresponds to progressive level of detail of the information on project costs and technical characteristics:

- the first tier includes general information on the technical characteristics of the projects (e.g. line category, geographical scope of the investment, etc.) and a preliminary categorisation of the total project costs into three sub-categories (planning & design, construction works and ancillary costs);
- the second tier focuses on two main parts of the rail infrastructure project - construction works (and associated costs) and ancillary costs.
- the third tier focuses on costs and technical characteristics relative to the equipment and permanent way, and tunnels, which are classified based on their construction method.

1.1.1.1 Validating the approach

The unit cost ranges identified analysing the total investment cost and the construction of rail projects clustered based on the type of works and infrastructure category, are reported in Chapter 5.

The total investment unit cost ranges per infrastructure category have then been broken down into smaller categories such as: constructions, base infrastructure and cost of single components. The good results achieved, in terms of difference between maximum and minimum values of the unit cost, confirmed the validity of the approach adopted for the analysis.
Indeed, the distance between maximum and minimum values of the unit total cost is lower than that that encountered in the majority of the previous studies (see Chapter 3), most likely as a result of the accurate classification activity carried out in building up the database.

Figure 1 – Breakdown of Total investment cost Tier 1

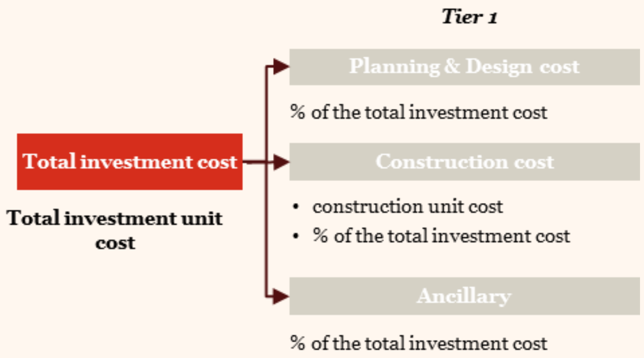
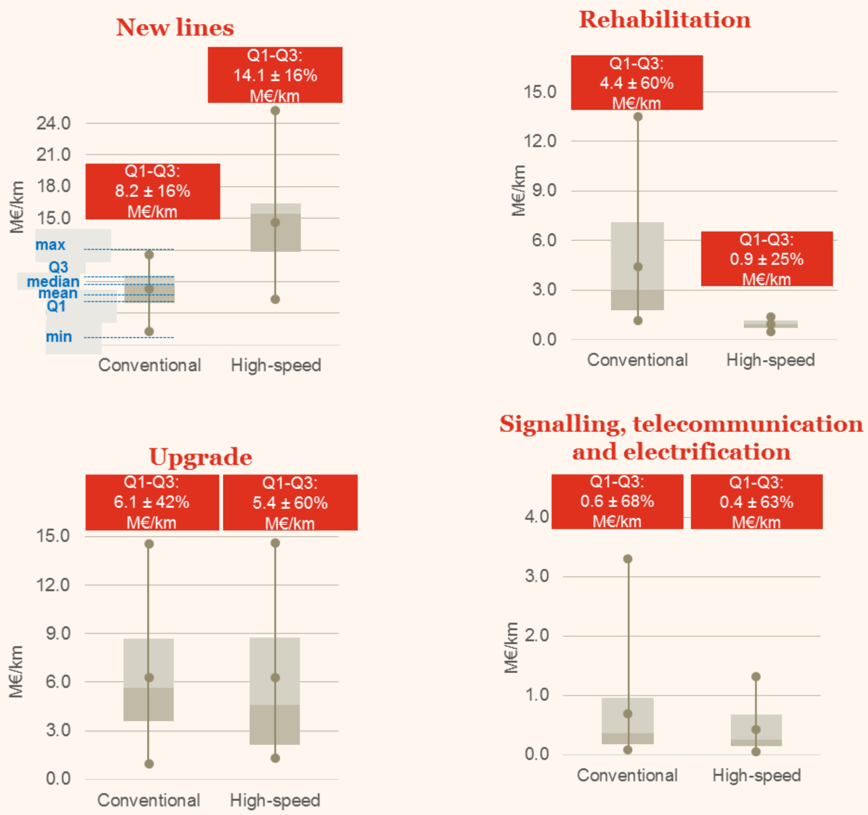
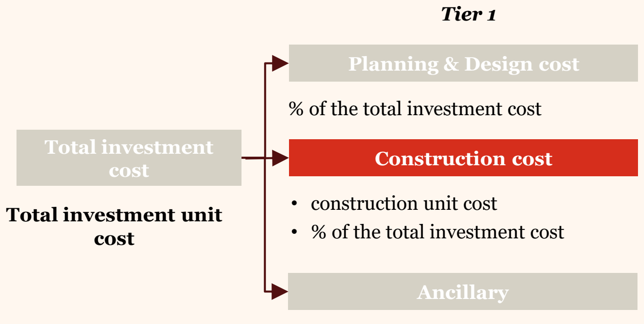


Figure 2 - Total investment unit cost ranges per infrastructure category



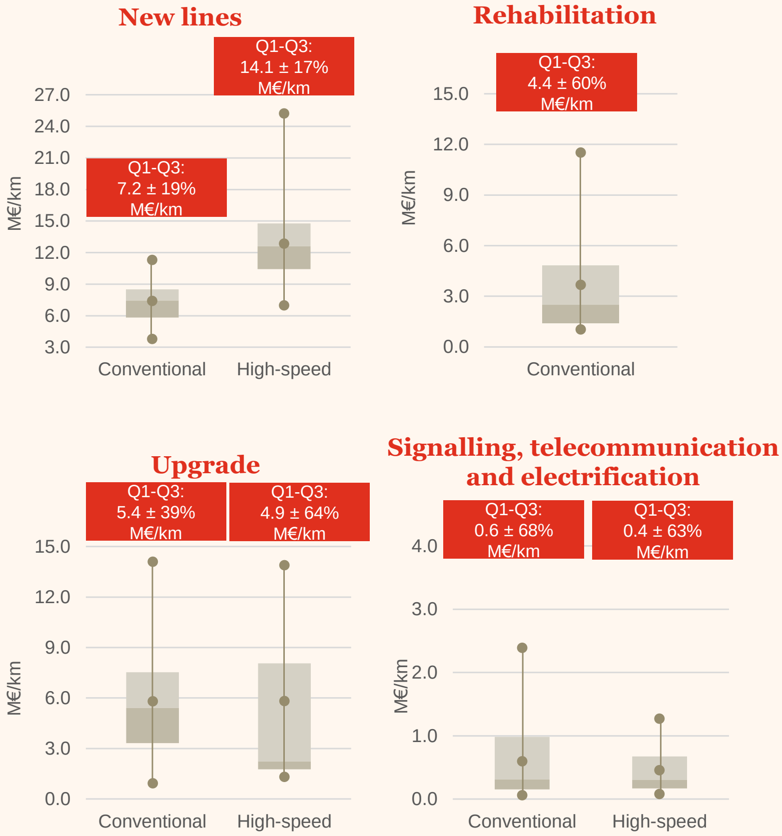
The analysis of the construction unit cost shows how, refining the scope of the analysis, more accurate cost ranges can be achieved, thus improving the reliability of the benchmarking and forecasting exercise.

Figure 3 - Construction cost Tier 1



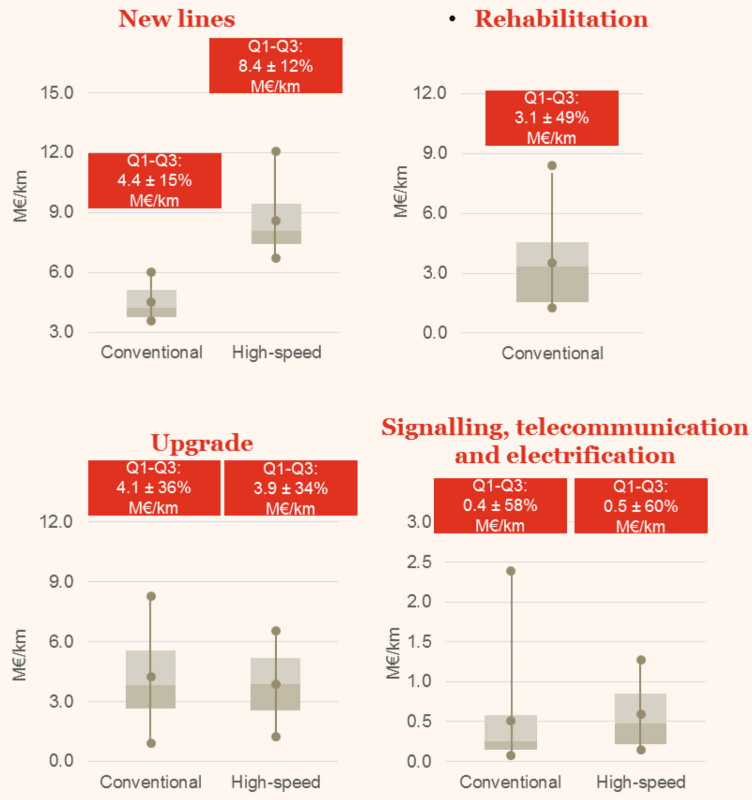
1.1.1.2

Figure 4 - Construction unit cost ranges per infrastructure category



The analysis of the base infrastructure unit cost enables to reduce overall the variability of the cost ranges, confirming that applying a bottom-up approach on an accurately clustered dataset leads to the identification of accurate unit cost ranges to be considered in a project comparison analysis.

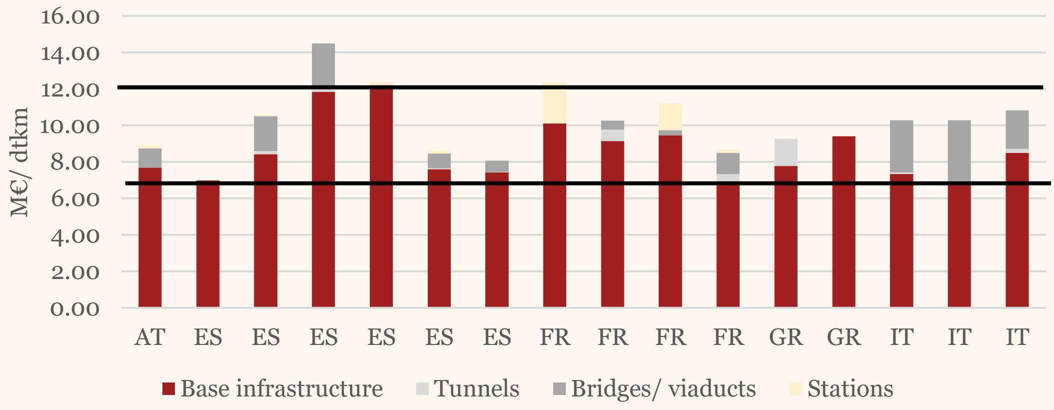
Figure 5 - Base infrastructure unit cost ranges per infrastructure category



Increasing the level of details, allowed to identify the different clusters and components, isolating the main cost drivers – tunnels, Bridges and stations – from the base infrastructure.

The following figure shows how, subtracting the cost of civil structures and stations to the construction cost of new high-speed lines, the unit cost ranges become closer.

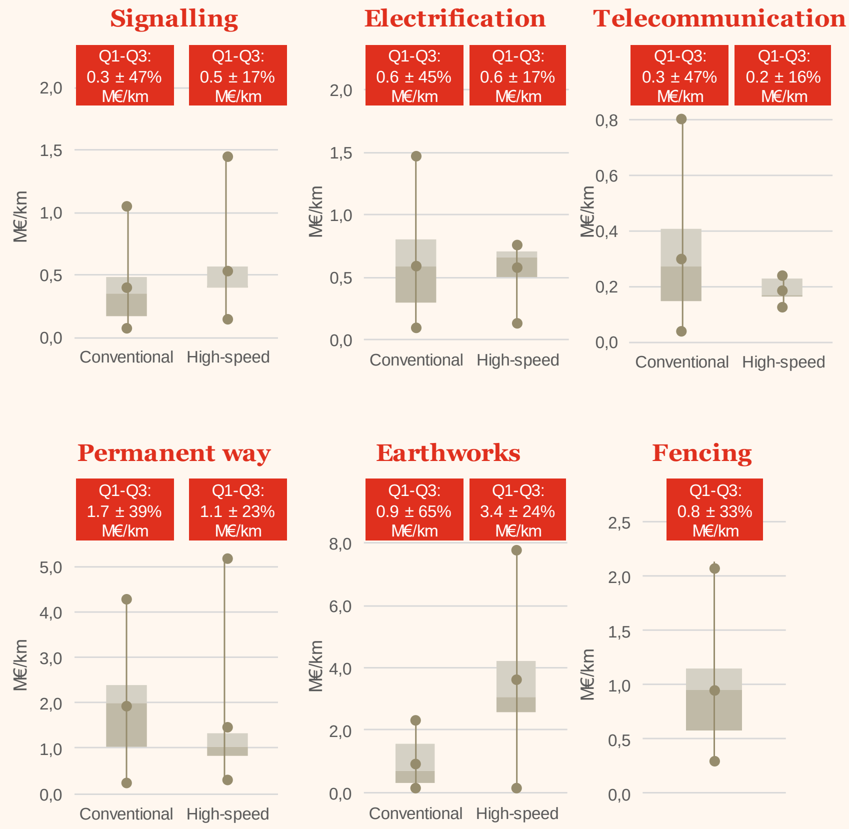
Figure 6 – Base infrastructure costs of new high-speed lines



It is worth mentioning that, concerning high-speed lines, the unit costs of the base infrastructure have been calculated on a sample representing approximately 45% of the high-speed railway lines in use in the EU.

The analysis of the single components to identify for each of them cost ranges that can be used to perform a reliable analysis.

Figure 7 - Unit cost ranges per single cost item



In addition to ranges of rail investment unit costs, the qualitative information gathered from the stakeholder consultations and by the experts of the Steering Group allowed to identify common patterns across Cohesion and non-Cohesion countries, in terms of elements leading differences between estimated and final costs. In particular:

- The higher level accuracy in planning process of non-cohesion Countries leads to limited differences between planned and final costs;
- In case of the implementation of new technologies, the lack of experience in estimating costs often creates discrepancies between panned and final costs;
- Different procurement practices and different marketing conditions create significant changes in final costs compared to the estimated ones.

Possibilities for further development

Despite being the result of a significant effort collecting and analysing data, and being designed to evolve over time, the REGIO Rail Unit Cost Tool cannot answer all the needs related to users assessing rail infrastructure works. It is indeed designed to respond to a specific objective. However, further opportunities for expansion have been identified.

The REGIO Rail Unit Cost Tool database was built with the information gathered from the Member States, the European Commission, INEA, JASPERS and Infrastructure Managers and it covered almost all the EU Member States with railway networks. As the REGIO Rail Unit Cost Tool results depends on the dimension and reliability of such database, this must be enriched over time, to maintain and increase the tool's functionality. Instead of dedicating resources to collect projects from all over the European Union once every few years, data can be collected constantly. To pursue this strategy, data collection shall be standardised in the way it can be input in the model. Furthermore, the REGIO Rail Unit Cost Tool may be shared with other entities as Public Authorities and Promoters, which can contribute to the database enlargement.

The tool is currently embedded in a Microsoft Office Excel file which represents a simple and user-friendly format, but does not allow a centralised management of the REGIO Rail Unit Cost Tool. It means that every time a file is shared, the whole database is duplicated and becomes a standalone tool. To overcome the trade-off between centralising the management of the REGIO Rail Unit Cost Tool and sharing it to collect as many data as possible, more advanced ICT architectures may be developed. A centralised server connected to e.g. a website interface may enable users from all Europe to assess their projects based on the same database.

Rail investments represent the most significant share of EU co-financed transport projects. Nonetheless, trans-European connectivity embraces other transport modes, as well. Motorways continue to represent the most used transport infrastructure. Ports are expanding to meet the increasing demand for goods from far ends of the World. Terminals to enable modal changes are more and more demanded by the logistics sector. As the economy and society develops, the need for infrastructure connecting us increases. Albeit covering a wide range of project investment types, this study focuses on a specific sector. Developing common understanding on how projects in other sectors are priced would produce benefits for all. It would be easier to assess projects; and also to support their planning. And it may hopefully ignite a debate among all actors on ways to e.g. better define cost categories, ranges, work types, etc. all contributing to ensuring an enriched and shared knowledge on the matter.

The REGIO Rail Unit Cost Tool provides an indication on the likely final cost of rail infrastructure investment by treating statistically the difference between estimated and final costs. Whereas, it provides no use understanding the specificities of single projects, which represent the risk factors impacting on costs. What it does, nonetheless, is facilitating action when specificities occur and reducing the range of elements to look closely at. The REGIO Rail Unit Cost Tool is not effective assessing risks. It is not the aim of the tool to assess the risks and analyse them into detail. Nor it could have been possible to develop such tool with the data currently available. Nonetheless, we are all aware of the significance of unexpected events on capital projects. Increased construction time, unforeseen geographical barriers, mistakes in or poor design, etc. are all elements that increase investment costs. And the probability that such events happen is usually relative to project-level specificities. Significant contribution to the present study would come from the identification, understanding and analysis of project risks. A risk-assessment-tailored study is deemed required to complement and complete the REGIO Rail Unit Cost Tool. Would the tool be upgraded with a risk-assessment add-on, it would enable its users to evaluate statistically the occurrence and severity of the impact of risks on the evaluated project.

1 Introduction

1.1 Background of the study

Bringing the people and economies of the individual Member States closer by developing a functioning, efficient and interconnected European transport system is a paramount objective of the European Union. [2] Together with activities related to regulatory oversight, the European Commission is therefore investing in transport infrastructure, aiming to create an interlinked network of European transport corridors. In this context, the development of the European rail network poses huge opportunities and currently represents a priority for the EU. The aim is in fact to increase the share of both freight and passenger transport by the rail, which represents the most environmentally friendly, sustainable mode of transport. [3]

In order to succeed in creating such a network, it is imperative to improve the efficiency and transparency of the public procurement procedures and planning of public funds at both EU and national level. Improving the process would be beneficial for many reasons including cost savings, a better selection process of the most projects representing best value for money as well as encouraging a more robust public private collaboration in future, and aid to all European organisations involved in the rail infrastructure planning by sharing experience and best practices with each other.

Previous studies have already provided beneficial insights into the public procurement process and unit costs of the rail infrastructure sector within and outside the EU. These studies, which have been used as a starting point for the present analysis (see 3.1), were generally focused on the transparency of the planning process and aid the relevant authorities to assess the cost of rail infrastructure projects more accurately.

PricewaterhouseCoopers (PwC) and COWI have supported **the European Commission, Directorate-General for Regional and Urban Policy** on the assessment of the unit costs (standard prices) of rail projects (Capital Expenditure).

Specifically, **the primary aim of the project** is to develop a tool[4] that can support the assessment of the investment cost of rail infrastructure projects, depending on project internal (e.g. technical standards) and external factors (e.g. environmental characteristics). To this end, the study assesses the costs of projects in the rail sector across 26 MS with rail systems and factors that lead to the variation of the costs of infrastructure investment in different MS, depending on different practices and external conditions.

[2] Among others:

European Commission 2016 Communication to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, *A European Strategy for Low-Emission Mobility* COM(2016) 501 final

European Commission 2016 Staff Working Document, *The implementation of the 2011 White Paper on Transport "Roadmap to a Single European Transport Area – towards a competitive and resource-efficient transport system" five years after its publication: achievements and challenges* SWD(2016) 226 final

European Commission 2011 WHITE PAPER Roadmap to a Single European Transport Area – Towards a competitive and resource efficient transport system COM(2011)0144 final

[3] European Commission 2011 WHITE PAPER Roadmap to a Single European Transport Area – Towards a competitive and resource efficient transport system COM(2011)0144 final

[4] The tool to be implemented at the end of the project is the DG REGIO Rail Unit Cost Tool. The REGIO Rail Unit Cost Tool will contain benchmark information of rail unit project costs differentiated into project types and components and information on the external country specific factors, which allows its user to evaluate a project based on its coherence with the standard unit cost.

According to the terms of reference, the main **objectives of the study** were:

- To establish common understanding on factors influencing rail infrastructure investment cost,
- To elaborate a methodological approach to identifying and calculating unit cost in rail infrastructure investments,
- To provide input to decision-makers for high-level evaluation of the costs of rail infrastructure programmes and projects,
- To assist authorities in charge of rail infrastructure programmes and projects to communicate effectively to a wider set of stakeholders about project cost.

1.2 Scope and limitations of the study

The scope of the present study lies on assessing the rail infrastructure unit cost in all European Member States with an existing rail network (the EU28 less Cyprus and Malta) and using the results to develop a tool to support the assessment of the investment cost of rail infrastructure project. The results of the study can assist public stakeholders involved in the procurement process to increase transparency and reduce costs in order to improve project planning and ensure optimal allocation of resources at project level. The analyses consider different types of works (e.g. construction of new tracks, upgrading, signalling and electrification) as well as different types of infrastructure elements (e.g. permanent way, equipment) and lastly different types of cost categories (planning, construction and ancillary).

In order to derive these insights, the technical and economic characteristics of circa 500 different rail infrastructure projects from all 26 Member States with a railway network have been analysed in detail. Further external factors which influence the cost have also been determined and taken into account. Unit cost ranges have been established based on past and ongoing projects, as well as discrepancies between planned and final costs, which can further be used to analyse the qualitative characteristics of the planning process in correlation to the cost.

The study provides useful insights on the rail unit cost, refining the results of the analyses previously carried out on this matter (see section 3.1). Nonetheless, these should be interpreted considering the following limitations:

- The majority of the data on the projects analysed was collected from DG REGIO and INEA databases, which comprised of the projects' applications to European funds supporting rail infrastructure investments (including TEN-T, CEF and ERDF) and from documentary review. Various stakeholders, in particular IM, have also been thoroughly involved in the process. The coherence of the data provided from all those sources was tested through data triangulation to the possible extent. However, it was not possible to perform due diligence on the information collected as the study is based on the availability of data provided by the stakeholders, which was considered to be reliable.

- It was not possible to ensure that the estimated cost data provided by the stakeholders and retrieved through the desk analysis refer to the same stage of the planning process (e.g. preliminary planning, detailed planning, feasibility study, etc.) and thus have the same level of accuracy.

- The level of detail reached by the analyses is a trade-off between the maximum detailing and the usability of the data. The calculation of statistics on specific cost-items enables to increase the accuracy of the analyses, yet the data analysed (and to analyse in the future) is provided by stakeholders. It cannot be too detailed, otherwise it could be only hardly collected, requiring a significant effort from the IM and/or the other organisations involved. This has been confirmed during the development of the study, when more detailed information has been provided for a lower number of projects (see section 4.4).

- Since the study is based on rail infrastructure projects from the European Union Member States for which data was available, its applicability for countries outside of those Member States cannot be guaranteed.
- The data collection process was originally designed to gather detailed information on both estimated and final cost of each project. Nonetheless, the information on the final cost of each cost item of the investment can be hardly retrieved, as the contracts for the project implementation do not usually refer to a single element of a project (e.g. provision and installing permanent way) but rather to a series of work to be implemented on a section of the line. Furthermore, the information on final cost is not frequently disclosed and cannot be collected from sources outside the IMs. It was quite informative to find out that IMs themselves are not free from interest politics and have strong inherent interests in the functioning and financing of the rail industry. For all these reasons, a limited amount of data on the detailed final cost of rail infrastructure projects was available (see section 4.4). The related analysis was therefore based on the total final cost, as outlined in section 5.
- In a few cases, IMs have been little active providing the information requested for the study. Reasons for this reluctance have been various, from the lack of resources available, to time constraints, to business-related reasons.

2 Capital expenditure in rail infrastructure

In the period covered by the study (2000 to present) the development of rail infrastructure across Europe was shaped by several regulatory and technological trends. Starting in 2001, the European Commission presented four legislation packages which focused on creating a single European railway network and market, while pursuing the economic development of the Union and implementing efficient, safe and environmentally friendly technologies.

Directive 2004/50/EC, updated by Directive 2008/57/EC and more recently Directive (EU) 2016/797, set out requirements for interoperability between the Member States' railway networks, thus allowing for the technical specifications to be harmonised among different countries and ensure that trains will be able to move all throughout the European Union on tracks with the same technical characteristics. This led to a heavy emphasis of European rail infrastructure expenditure on adapting national rail networks to the unified EU standards (e.g. the Standard Gauge width of 1435 mm) particularly in areas outside of Central Europe (e.g. Spain, Ireland, Portugal, and Central and Eastern European countries previously belonging to the Warsaw Pact).

In line with this objective, the European Commission has promoted and invested into cross-border projects between Member States. Specifically, countries formerly part of the Warsaw Pact which were not well-connected to the European Union, but whose rail network was Russia centric, invested strongly in upgrading the rail network and increasing connectivity with Central European countries. Furthermore, many of the Eastern European countries conducted works to upgrade their rail lines from single to double tracks.

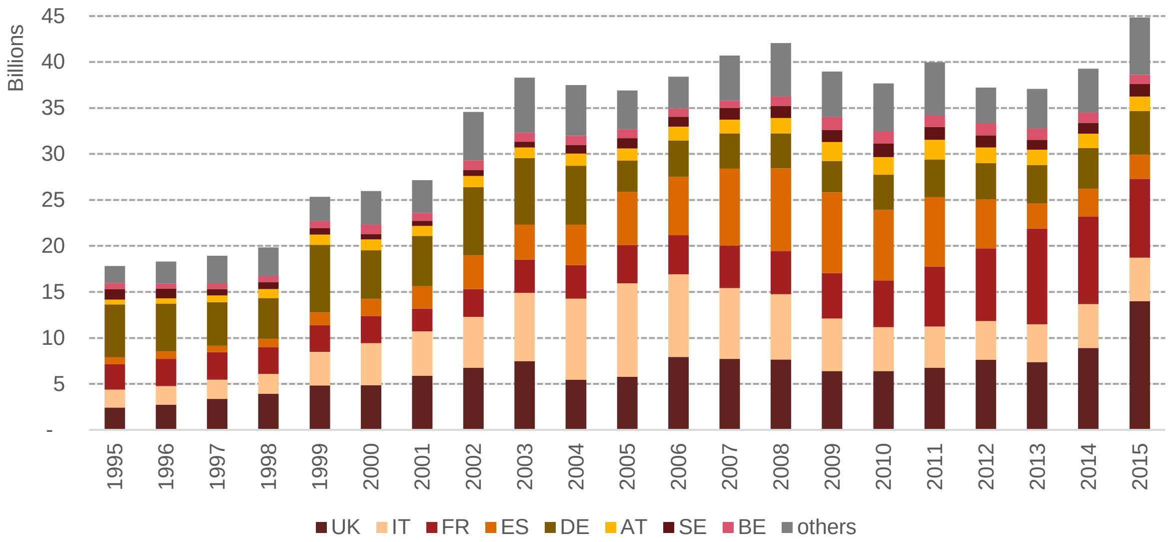
Additionally, the introduction of the European Railway Traffic Management System (ERTMS), which requires all Member States to switch from a national to a European signalling system, has shaped the infrastructure investments undertaken. All MS adopted this as the standard signalling equipment for new lines belonging to the core network and MS with well-established existing rail networks started to exchange the previous signalling equipment to adhere to the new requirements.

Besides the direction set out by the European Commission, technological advances on the field have led to an additional infrastructure investment focus on the development of a high-speed railway network, predominantly in Western European countries such as France, Italy, Spain and Germany. In other countries high-speed railways have been – and are being – developed as well, but on a smaller scale.

All of these elements and factors contributed to a significant level of investments in rail over the period from 2000 until 2015, which in the entire EU equals a total of almost €600 billion.[5] The following graph shows the total rail infrastructure investments over the period for the entire EU and the eight most investment-intensive countries.

[5] PwC elaborations on Data from OECD

Figure 8 - Rail Infrastructure Investment (in Billion €)



Source: PwC elaboration on OECD data

Despite increasing investments over time, the economic crisis in 2008 represented a turning point in the railway sector: the public sector had less financial resources available to invest and therefore made cuts on infrastructure development. Albeit the support put in place from the European institutions – of approximately €32 billion for infrastructure investment in 2009-2010 – which was introduced to partially sustain capital investment for a short time, railway development programmes stalled or were postponed. [6] Therefore, it took seven years to reach the same investment levels as in 2008.

In fact, the Member States which have invested the most in railway infrastructure development, are those that have undertaken a high-speed infrastructure development programme.[7] The figure above clearly shows that the countries with the highest investment levels in rail infrastructure correspond to the countries that undertook major high-speed investment programmes in the respective years.[8]

In contrast to the development of high-speed networks in most rail-advanced Member States, a significant investment was made to modernise and renovate the railway system in countries where the network was underdeveloped and obsolete.[9] Indeed, the Cohesion Policy investments in the 2007-2013 programming period resulted in the upgrading of ca. 1,600 km of railways and in the building of only 22 km of new railway lines. Several Cohesion countries - especially in Central and Eastern European Countries – present an underdeveloped railway infrastructure system, which required (and still requires) significant investments to reach the quality and efficiency standards of the Union.[10]

The figure below presents the wide discrepancy among Member States that developed a rather efficient railway transport system (mostly non-cohesion MSs) and those that have to significantly upgrade their

---

[6] among others, http://ec.europa.eu/economy_finance/publications/occasional_paper/2014/pdf/ocp203_summary_en.pdf

[7] The Netherlands, which invested almost as much as Belgium, has been recently investing in high-speed as well.

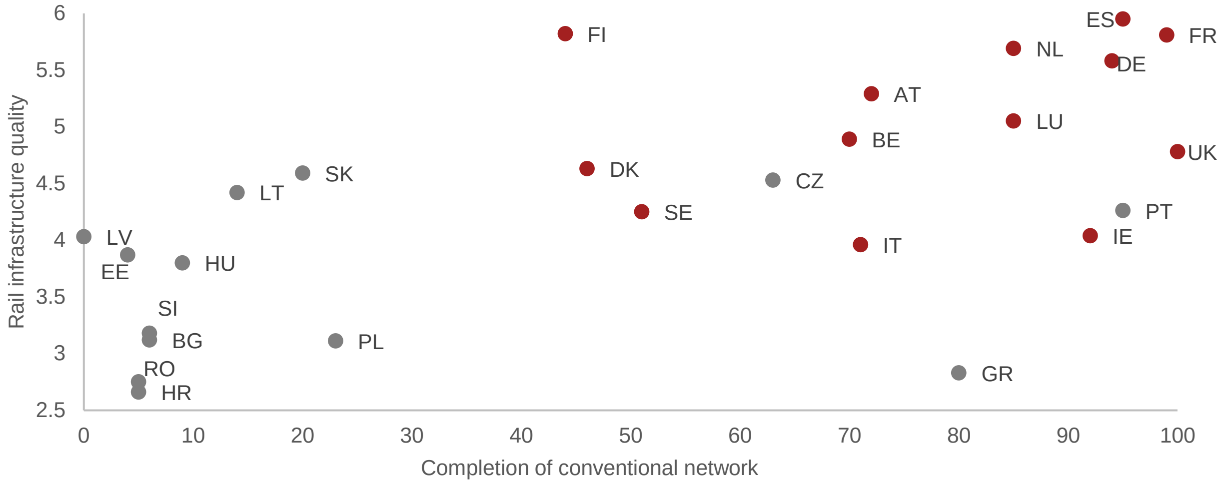
[8] I.e. in Italy between the late 90s and 2009; in Spain during the mid-00s and the early 10s; while France has a longer history in high-speed lines, recent investments occurred in the southern part for the SEA from 2011, the Rhine-Rhone in 2011, la LGV Mediterranée, etc.

[9] Ex post evaluation of the ERDF and Cohesion Fund 2007-13, EC, 2016

[10] Among others, World Economic Forum (see Figure below); European Commission, DG IPOL, Policy Department B: Structural and Cohesion Policies 2016 Research for TRAN committee: Connectivity and Accessibility of Transport Infrastructure in Central and Eastern European EU Member States (available at: http://www.europarl.europa.eu/RegData/etudes/IDAN/2016/573419/IPOL_IDA(2016)573419_EN.pdf)

systems to keep pace with both the market standards as well as with the EU requirements set out in the TEN regulations, among others.

Figure 9 - Level of development of the railway network in the EU per MS



Legend: in red the non-cohesion MSs; in grey the cohesion MSs.

Source: The quality of the railway infrastructure is derived from the World Economic Forum, 2016 World Economic Forum Global Competitiveness Report; the level of completion of the conventional network considers the portion of completion of the conventional TEN-T Core network as by European Commission TEN-Tec.

The level of development of the railway infrastructure system does not only relate to the length and quality standards of the lines. It also represents a proxy of the degree of development of the market that generates from rail investments, e.g. railway constructing companies, technology providers, engineering companies, etc. In some countries, only very few companies exist that are able to carry out rail infrastructure projects due to their complex nature. This often leads to low competition in the construction market, resulting in final costs higher than estimated ones, as confirmed by ca. 25% of stakeholder responses involved in the development of this study.

Nonetheless, the response to the crisis both in terms of the market adapting to the mutated conditions[11] and the public sector simplifying procurement procedures enabled to partially overcome the consequences of low competition. Indeed, in the years after 2008, the majority of the stakeholders[12] reported that higher competition in the construction market resulted in a decrease of final investment cost, contrary to the pre-2008 period. Furthermore, the level of competition in the construction market varies strongly among the different countries e.g. in some countries a high amount of single bids is received for tenders (e.g. Poland with 46% single bidders), which as a consequence can have an impact on the cost of the contracts.

Member States are also quite different in terms of procurement processes. Accordingly, different effects on prices occur. In several cases, the different awarding criteria across countries have shown a large influence on the final project prices. Particularly in Cohesion Member States the awarding criteria were often based on the lowest prices. In many cases this requirement was subsequently revised - particularly because of the introduction of the recent European Directive on public procurement[13] - to ensure that the quality of the offers were awarded a score as well. Nonetheless, the use of quality criteria varies between Member States,

[11] Companies were more strongly competing for international projects due to reduced spending and investments. Additionally, particularly after 2008 in several countries a drop in material and labour prices benefited lower overall prices.

[12] Over 65% of the respondents to the qualitative survey which identified the competition in the construction market as a relevant factor in the difference between estimated and final costs.

[13] Directive 2014/24/EU of the European Parliament and of the Council of 26 February 2014 on public procurement and repealing Directive 2004/18/EC Text with EEA relevance

from below 10% in Croatia and Lithuania to over 90% in France and the United Kingdom. [14] In general, many Member States in Central and Eastern Europe still use primarily the lowest price criterion.[15] This has often led to cost savings in many countries due to bidders undercutting prices of original estimates. Furthermore, several other indicators of the efficiency of the procurement process - e.g. administrative capacity, length of the procurement procedures, level of transparency – that vary from country to country impact the difference between estimated and final cost of rail infrastructure projects.

Additionally, the planning process of the infrastructure projects varies widely across the different Member States and can lead to significant differences in cost of the railway expenditures. Particularly, in applying for co-funding programmes, some MS (mainly occurring in Cohesion Countries) estimate the cost for projects either only in general detail or as a lump sum, which might lead to larger discrepancies between the estimated and final costs, while other Member States use detailed estimates at each stage of the planning phase (e.g. Austria).

Furthermore, other characteristics and trends about the European rail infrastructure investments over the analysed period can be observed. Among the legislative factors particular technical standards had a large influence on the cost of rail infrastructure projects. These standards relate mainly to the Technical Specifications for Interoperability (TSI), safety related standards and requirements for the ERTMS. While these standards are set at EU level, there are differences across Member States, which in some cases require even more stringent rules or are characterised by conditions that make it more difficult and expensive to comply with specific legislative requirements. For example in mountainous regions (e.g. Austria, Northern Italy, Southern Germany), tunnel safety standards can drive costs up. Additionally, environmental regulations regarding protected territories and building sites, which are part of EU law, had a strong influence on the rail infrastructure cost. For example in the Czech Republic, the implementation of the EU laws on environmental protection had a strong effect on the cost of infrastructure projects. Lastly, in some Member States complicated legal procedures for the purchase of land for new railway lines drive the final cost of the projects.

The analysis of the rail infrastructure investments clearly outlines the differences that can be encountered among the European MS. The objective of this study is to create **a benchmark that has validity on a European level**, therefore, a tailored approach to normalise the differences between the economies of the European MS has been adopted (see section 4.3.1).

The normalisation enables to take into account the differences encountered in the analysis of investments carried out in different countries and in different time periods.

It does not embrace all the specific aspects of the rail infrastructure industry that impact investment cost (e.g. level of detail of the planning process, length of the procurement procedures, level of competition in the construction market, implementation of interoperability and safety requirements among other), which vary from country to country.

These have been analysed during the stakeholder consultation and the literature review and compiled in **individual country sheets for each Member State** that complement the results of the quantitative analysis (see Annex 11).

[14] Single market integration and competitiveness report 2016, EC, 2016.

[15] *Ibidem*

3 Approach to the study

Building on the learnings that can be derived from the previous studies, this study was designed to circumvent the difficulties and limitations that authors of similar analyses encountered. At the same time the methodologies that have proven to yield successful results have been extensively adopted.

Hereunder, the main studies and publications that have been analysed are reported. For a complete list of documentation used to develop the study, do refer to the bibliography in annex.

3.1 Building up from previous studies

Approaching the study, it is important to make best use of the results that others reached on the subject. This way, the lessons learned as well as the relevant outcomes can be incorporated without replicating the effort. At the same time, the experience of the past can suggest approaches to undertake or to abandon, based on the authors' explicit experience. Therefore, the initial step of the study consisted of the analysis of existing studies on the field of analysing rail infrastructure projects on their cost. This activity aims to understand what has been done so far, how it has been done and which results have been achieved in order to use those insights to better approach this current study.

While analysing the studies, two main limitations manifested themselves. The first one is the difficulty to gather a comprehensive and complete set of data for the analysis, meaning that, overall, the studies analysed tended to focus on a limited amount of data, which was available for the analysis. Although the studies showed valid results, often the data set was too small to generalise the findings on the (European) rail infrastructure environment as a whole. Secondly, the majority of studies had a small technical scope, most likely related to the amount of data available. Therefore, unit costs were provided either on a whole project basis or broken down into highly approximate rough subcategories, which limits the applicability of the study to improve detailed planning of rail infrastructure. To this regard, it is worth mentioning that studies concerning the analysis and benchmark of infrastructure costs may be carried out with different purposes (e.g. supporting the analyses of potential cost-reductions in infrastructure projects and programmes, increasing the effectiveness of communication to the public, analysing the potential for cost-reduction, etc.). The studies analysed do not necessarily share the same objectives of the present one (as detailed in Table 4).

3.1.1.1 The objective of the benchmarking analysis

Benchmarking exercises differ depending on the objective of the studies they are developed within and so does the methodological approach to carry them out (e.g. in terms of data disaggregation level; sample for the analysis; etc.). In the framework of this study, the most relevant objectives to achieve include:

- providing input to decision-makers for high level evaluation of the costs of infrastructure programmes and projects;
- supporting the analyses of potential cost-reductions in infrastructure projects and programmes;
- assisting authorities in charge of infrastructure programmes and projects to communicate effectively to the public opinion about project costs and related drivers.

Despite such limitations the studies analysed provided very insightful results, which were used to refine our approach in order to on one hand not encounter the same limitations and on the other hand use the results to investigate certain findings in more detail. Additionally, one of the main benefits the studies provided were their methodological approaches, which proved to be very valuable to optimise the methodology followed in this study.

In the following pages an overview of the most relevant studies, their results and limitations as well as a comparison among them is composed.

Baumgartner

Baumgartner's study is one of the most renown in the railway sector regarding costs of both infrastructure and rolling stock. Although it was published already more than 15 years ago (2001) it is still today often quoted for railway sector assessments, because it provides a fairly detailed assessment of costs for a wide range of elements covering nearly all aspects of the rail network (rails, trains, tunnels, bridges, stations, and maintenance for these elements).

Baumgartner specifies the cost analysis to be based on providing the average of median value and two extreme values. Nonetheless, the exercise from Baumgartner is by its author's own admission a very broad one and he disclaims from providing a price list and advises users not to rely on this document for creating financial or economic studies.

While the study is very interesting for a methodological approach, in the way structures are broken-down into sub-clusters, it was not possible to benefit from the experience on actual projects used, nor on the outcomes. Indeed, it is not specified how the data to derive these cost measures is gathered and calculated. Additionally, the absolute cost ranges of the projects are not specified, as Baumgartner stated that the "limits in brackets do not take account of special or exceptional cases".

The age of the paper and the costs it was referring to (Baumgartner specified that the costs were reflective of the year 2000) further leads to a decrease in accuracy for recent and future rail infrastructure projects.

Moreover, Baumgartner applied a top-down approach. The breakdown into individual type of works on the tracks has been carried out very generally. The cost ranges specified are inclusive of many elements (such as work management, rerouting of roads, tunnels, bridges and other structures etc.), whose breakdown is not further elaborated on. While, it is not possible to carry out accurate estimates based on the proposed figures, the general breakdown and considerations were taken into account in designing the tool for data gathering (see Annex 3). Similarly, aspects impacting on rail costs that are suggested by Baumgartner are taken on-board to ideally identify their impact, if possible to a higher level of detail than in his study. Baumgartner's analysis also took into account the topography in which the line is built and differentiated costs accordingly, but did not specify the definition of the topography parameters.

The main results – unit cost of railway line, bare tunnels and bridges, are presented in the following tables.

Table 1 - Baumgartner's results for rail unit cost of different line types in different topographies (inclusive of all cost elements)

Type of Track	Easy Topography Avg. (range) in M€/km	Average Topography Avg. (range) in M€/km	Difficult Topography Avg. (range) in M€/km
Single 100 km/h	2 (1-3)	5 (3-15)	20 (15-40)
Double 100 km/h	2 (1-4)	7 (3-20)	20 (20-50)
Double 300 km/h	3 (2-6)	10 (6-30)	40 (20-50)

Table 2 - Baumgartner's results for rail unit cost on bare tunnels (excluding rail infrastructure)

Type	Avg. (range) in M€/km
Single Track	20 (10-50)
Double Track	30 (20-70)

Table 3 - Baumgartner's results for rail unit cost on bridges

Type	Avg. (range) in M€/km
Short Span/Easy Foundation	15 (10-20)
Long Span/Difficult Foundation	30 (20-50)

Technical University of Denmark

The paper of the DTU (Danmarks Tekniske Universitet) is focused on benchmarking European high-speed railway projects to create a procurement guideline for the first Danish high-speed line construction (Copenhagen-Ringsted) by comparing its characteristics to similar European projects. It focuses particularly on avoiding cost overruns by analysing historic cost overruns in comparable rail projects.

For this purpose the study analysed 19 different European rail infrastructure projects of mainly high-speed rail construction, with few exceptions of conventional line construction. The projects came mainly from countries in Central and Northern Europe with the exception of three Italian and one Spanish project. The analysed projects were purposefully selected to be from countries with similar conditions and a shared construction market, such as the Netherlands, Belgium, Germany, Sweden and Norway. This allowed the authors of the study to benchmark the Danish project against projects more similar to it than projects from an international or entirely European landscape.

Evidently, by nature of this analysis its results are focussed on rail development cost observed in a specific region. Additionally, the analysis concerns high-speed structures only and can therefore provide useful hints to assess alike projects.

Unfortunately, as the authors of the study specified, it was very difficult to obtain the complete and detailed information from the stakeholders (in particular Infrastructure Managers and Construction Companies) because the data was either confidential or not accessible.

As a result, information on detailed unit costs based on breaking down the cost into individual elements could only be provided for three projects (including the Danish project), as not sufficient data was available for the other projects. Additionally, the study provided information for the cost per kilometre of track construction, earth works and civil works, as well as signalling and land acquisition. This information was provided only in an anonymous format, wherefore it has not been possible to use the data. Nonetheless, the study's approach to differentiate costs into homogeneous categories has reassured the approach used in this present study, as well as the categories that were used.

Out of the original 19 projects analysed, the authors specified to have been able to analyse ten average costs per kilometre ranging from 11.47 M €/km to 62.72 M €/km. Moreover, the study stated that the cost range for the railway track construction (as an isolated element) per kilometre went from around 1 M €/km to nearly 5 M €/km with an average of 2.35 M €/km. For eight projects a cost range from slightly above 2 M €/km up to ca 16.5 M €/km was provided for earthworks (average 6.3 M €/km). For civil works the cost ranged for ten projects from 2 M €/km until nearly 14.5 M €/km with an average of 6.82 M €/km. Data on signalling and land acquisition varies strongly because of regional differences in project focus and laws regarding land acquisition.

Overall, the study achieved good result for its purpose which was to benchmark the overall cost of the project per kilometre against other European projects. It provides a lot of potential scope to build on for this present study in terms of the European rail unit cost as a whole and per region or Member State. Nonetheless, the limited sample used does not enable to derive statistically significant findings, and detailed enough information for individual rail infrastructure elements is not provided.

PwC Study on HS2

Similarly to the Danish study mentioned above, the study conducted by PwC was carried out with the purpose to benchmark the British "HS2 Phase Two capital costs against other international high speed rail projects."[16] This exercise further intended to identify certain opportunities for the project, which could be used to reduce the costs or duration of the construction.

The study compared 20 European high-speed projects that are used as comparators and standardised by normalising the prices to a £ equivalent. PwC placed particular focus on assessing similarities and differences between the analysed project and HS2 Phase Two by evaluating different topographical characteristics, ancillary cost, indirect cost, land and property as well as contingency costs, and lastly civil engineering buildings such as tunnels and viaducts.

The costs of the projects range from 12.5 M €/km to 89.6 M €/km with an average of 36.3 M €/km. Part of the project clustering by topological factors resulted in the finding that sections of comparator projects that run predominantly through rural environment lie at the cheaper end of the spectrum with a cost range from 12.5 to 22.7 M €/km. Urban project sections and sections in an area of high infrastructure density make up the upper-medium spectrum due to the need for having frequent over- or under passes; these structures ranged from 48.8 to 69.2 M €/km. Sections with complex assets structures such as predominantly running underground through tunnels made up the high end of the spectrum with a range of 81.7 to 105.5 M €/km[17].

Additionally, the study further elaborated on a breakdown of the ancillary costs per kilometre or project type. Cost elements such as design, insurance and project management were considered as indirect costs and combined made up between 7.5% and 20.4% of the total construction cost (measured on eight comparator projects). Land and property costs for the land on which the infrastructures were built ranged from 0.5 to 6.4 M £/km (measured on six projects). Contingency allowances against unforeseen or uncertain cost elements were said to have accounted for 10%-30% of three projects that were at similar stages as HS2 Phase Two at the time of the study.

Lastly, the study provided insights on the cost of civil assets, such as tunnels, viaducts and earthworks. Tunnels of the comparator projects ranged from 20.4 to 70.3 M £/km and had an average cost of 40.8 M £/km. This was observed on seven projects. For viaducts, which were also observed on seven projects, the range was from 14.7 to 60.1 M £/km and a very similar average to the tunnel structures of 35.1 M £/km. For earthworks of four comparator projects, the range went from 4.5 to 9.1 M £/km with an average of 6.8 M £/km.[18]

The methodology followed in the study was taken into consideration in the development of the present analysis. Indeed, a similar bottom-up approach has been followed to derive unit cost (see 5.1). Additionally, the report provided information on costs and technical specifications of ten projects which, after being integrated and cleaned, was included in the database used in the analysis.

Other Studies with Rail Unit Cost Elements

Besides the abovementioned studies, which were the most similar compared to the scope of this present study, several others have been reviewed, whose characteristics are summarised below.

In 2007, the Indian University IRICEN published the report "Case Studies – High Speed Rail System (> 250 kmph)" in form of a collection of case studies of different European and Asian rail projects to provide an initial guidance document for the planning of an Indian high-speed rail network. It uses twelve of these projects and demonstrates the characteristics of high-speed rail lines and networks by providing a comparison of different technical parameters and unit costs. The unit costs of the twelve projects range from ca. 10 M €/km to ca. 70 M €/km. The report's focus was not to provide accurate unit costs for the

[16] (PwC, 2016)

[17] Convertion rate 1 GBP = 1.13449 EUR as per 03/05/2018
[18] ibidem

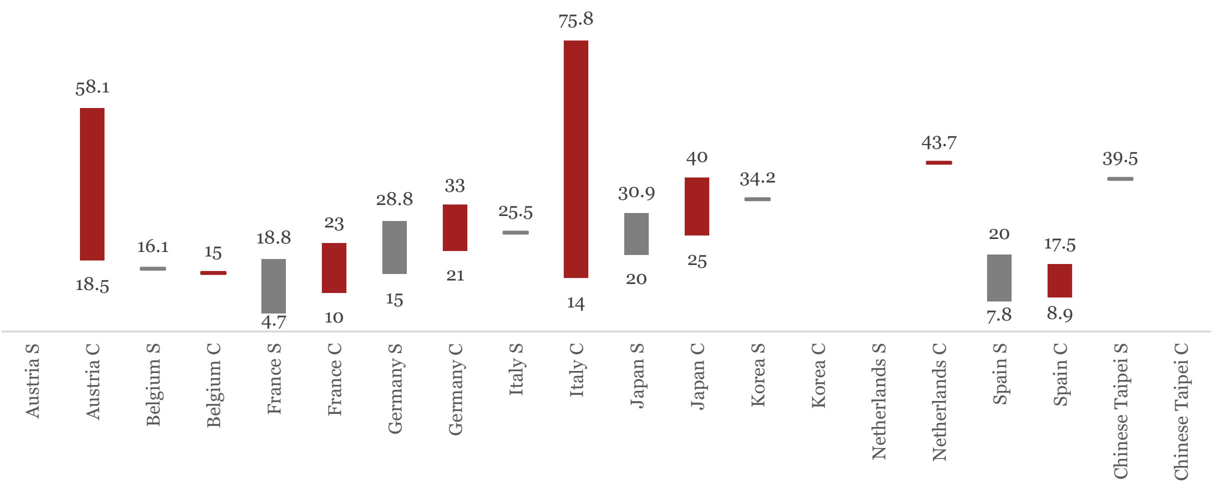
planning of infrastructure projects, but to provide an international overview of the high-speed rail landscape.

The Final Report of Work Package 10 of the Ex Post Evaluation of Cohesion Policy Programmes from 2009 focused on assessing cost overruns, delays and unit costs of different types of infrastructure projects, including rail. For rail in particular 24 projects were analysed, nonetheless the analysis of unit cost could be carried out for a total of 20 projects due to a partial lack of data of the other projects. The rail cost was analysed on two levels (overall per kilometre and construction cost only per kilometre). The total costs per kilometre range from less than 1 M € to 90 M €. Additionally, results are presented in different ways in the different sections. Nonetheless, the study is based on a solid and thorough methodology, providing useful insights for the present analysis. Specifically, the same criterion for the organisation of information on costs and technical specifications – which was based on progressive level of detail - has been adopted and followed in the definition of unit cost ranges (see section 5.1).

For the same context, the Work Package A presented a Second Interim Report in 2011 in which it analysed different cohesion projects of which 82 were rail infrastructure projects. These were all projects co-financed by the Cohesion Fund and ISPA in the 2000-2006 programme period. The unit cost ranges for those rail projects were again split in Level 1 and Level 2 details with a total cost range of actual projects costs from 1.22 M €/km to 111.9 M €/km.

In 2009, the Fundación BBVA published an "Economic Analysis of High Speed Rail in Europe" which focuses on providing insights to governments on the advantages and benefits of high-speed rail infrastructure investments. During their analysis, the foundation elaborated on UIC's HSR Database to derive average costs of high-speed rail infrastructure in different countries. The data is based on information from the years 2005 and 2006. The average cost ranges per country split into lines already in service and still under construction in 2006, as shown in the table below.

Figure 10 - BBVA cost ranges per country for lines in service (S) and under construction (C)



Legend: in grey, the countries' average cost range for high-speed lines already in service in 2006; in red, the average cost ranges per country for high-speed lines under construction in 2006.

Source: Economic Analysis of High Speed Rail in Europe, BBVA, 2009

The report further specifies that in France and Spain the costs are lower than for the rest of European countries, because of less densely populated areas and geographic reasons. Additionally, in France steeper grades are used instead of using a high number of civil engineering structures (tunnels, bridges, etc.) permitted by segregation of passenger and freight services (a mixed line requires higher axial load bearing capacity than a dedicated line for passenger services alone). The large cost range in the projects under construction and their discrepancy with the line already in operation in Italy is explained because of a higher density of population in the North than in the South of the country.

Another report from Albalate and Bel “The Economics and Politics of High-Speed Rail: Lessons from Experiences Abroad” provides an overview of the development of high-speed rail lines using several international case studies. Furthermore, it provides certain insights on how the unit cost has changed over time. For France in particular it provides the examples of the lines Paris-Lyon (1981 - $11 Million per mile), Méditerranée (2001 - $42 Million per mile), Est-partial (2007 - $27 Million per mile), and Rhin Rhone-partial (2011 - $52 Million per mile).

In the following table the studies above and other insightful documents useful for the analysis are summarised to allow for an easy comparison.

Table 4 - Summary of the previous study analysed

Study	Rail Focus	Time (Cost Data)	Technical Range of Analysis (Projects)	Geographic Scope	Outcomes (Unit Cost Ranges in M€/km)	Purpose of the study
Baumgartner	Rail (HS & Conv.)	2000	Unit cost is provided for very detailed infrastructure levels (e.g. track only, tunnels, earthworks, signalling, electrification, bridges	US and Europe (no specific countries mentioned but the prices were provided for Europe and the US)	Single Track 1-40 Double Track 1-50 Tunnel (only) 10-70 Bridge (only) 10-50	Support the decision makers in identifying reference prices of equipment and infrastructure
Technical University of Denmark	High-speed Rail	1993-2025	19 projects were analysed of which for ten the unit cost could be calculated; mainly high-speed lines with a focus on benchmarking the Danish line against others Mainly total cost analysed, but also Civil Works, Earth Works, Signalling, Electrification	BE, DE, DK, ES, FR, IT, NL, NO, SW, UK,	Double 11.47-62.72 Civil Works 2-14.5	Identify of causes of cost overrun
PwC Study on HS2	High-speed Rail	2011 (norm-alised)	20 cost comparator projects; topographical characteristics, ancillary costs, indirect costs, land and property, contingencies, tunnels and viaducts	AU, BE, CN, DE, ES, FR, IT, NL, NO, TW, UK, US	Double 11 – 79 (M£) Tunnel 18-62 (M£) Viaduct 13-53 (M£)	Identify potential opportunities for reducing the costs or duration of the construction
Indian University IRICEN	High-speed Rail	Un-known <2007	The report provides an international outlook on high-speed unit cost (total project kilometres) on 12 projects but does not enter into any technical detail	DE, ES, FR, IT, JP, KR, NL, TW, UK	Cost per Project km: 10 - 71	Provide an international overview of the high-speed rail landscape

Study	Rail Focus	Time (Cost Data)	Technical Range of Analysis (Projects)	Geographic Scope	Outcomes (Unit Cost Ranges in M€/km)	Purpose of the study
Final Report WP10	Infra-structure	2000-2006	The analysis seems to focus on different types of rail projects (20 projects for which cost is analysed), but it is not specified which technical specifications they contained The report is not focusing on rail infrastructure only, covering also other types of transport and energy infrastructure	DE, ES, GR, IT	Convent. DT 20-90 High-speed DT 3-12	• Identify of causes of cost overrun; • Assist authorities in charge of infrastructure programmes in the appraisal of future financing requests
2nd Interim Report WP A	Infra-structure	2000-2006	Different types of rail projects (total 82) are covered and the detail is split into Level 1 and Level 2 (overall cost per km and built cost only) exact technical specifications are not explained The report is not focusing on rail infrastructure only, covering also other types of transport and energy infrastructure	BG, CY, CZ, EE, ES, GR, HR, HU, IE, LT, LV, MT, PL, PT, RO, SI, SK,	Cost per Project km: 1.22-111.9 Median 7.23 St.Dev 16.11	Assess the efficiency of infrastructure projects co-financed by European funds
Report on Bridges for High-speed Railways	High-speed Bridges	No Cost Data	The report provides significant insight on the technical details on high-speed rail viaducts; cost is not a primary focus, but the technical details allow for an insight in cost driving factors for specific case studies of this current study	Europe, in particular: ES, FR, IT, NL Asian seismic bridges	N/A	N/A (not a benchmarking study)
Boletín Oficial del Estado	High-speed Rail	2010 (base year)	The Spanish State bulletin is a guideline for cost planning and comprises of different high-speed rail infrastructure elements, such as the line in different types of terrain, electrification and viaducts	Spain	Double 2-16 Electrification 1.1-1.35 Signalling 1-1.25 Viaducts 800-2500 (€/sqm)	Provide general statistics on costs per infrastructure level to support the planning.

Study	Rail Focus	Time (Cost Data)	Technical Range of Analysis (Projects)	Geographic Scope	Outcomes (Unit Cost Ranges in M€/km)	Purpose of the study
Fundación BBVA: Economic Analysis of High Speed Rail in Europe	High-speed Rail	2005	Unit cost is provided on a cost per kilometre on project basis elaborated out of the UIC database	AT, BE, DE, ES, FR, IT, JP, KR, NL, TW	Double 4.7-65.8	Assess economic benefits and costs of investments in high-speed railway lines

The analysis of the most recent literature on the matter demonstrates that a relevant variability exists on the unit costs of rail infrastructure, as reported by all studies. This variability limits the usability of the results in practical cases. It is therefore the purpose of the approach followed in this study to identify methodologies to explain such variability and, whenever feasible, reduce it by breaking down the cost into sub-components.

Further, a lack of uniformity and detailed information seems to have impacted on the assessments made on wide scale databases, e.g. the Cohesion Policy reports. In these cases, a statistical approach was applied on non-treated, non-revised data, and eventually resulted in significant variability. Oppositely, the researches built on a limited sample of cases, which have been analysed in higher detail (e.g. TUD, HS2) presented less dispersed results, albeit still far from being within a useful range for the purpose this study aims to – which is, however, different from the specific purpose of the studies reviewed.

Therefore, the aim of this study approach is to overcome these issues by analysing a large quantity of data to the highest possible level of detail.

3.2 Our approach

In line with the examples provided by the previous authors challenging the subject of identifying common unit costs in railway infrastructure projects, our approach was primarily based on:

1. Collecting the data necessary for elaborations to be performed. The analysis of previous studies highlighted how the limited amount of data available was one of the causes impacting the accuracy of the analysis of the unit cost of rail infrastructure investments - increasing their variability. Therefore, a significant effort was made to collect a large amount of projects to build the analysis on. Specifically, as outlined in section 4.2, a number of rail experts and dedicated teams operating on-site have been involved in the analysis of project documentation and in retrieving information directly from the IM over Europe, for a **total effort of over 450 man/days**.

   Annexes 1 to 6 present the different activities and background documentation used in the data collection phase, as well as the list of organisations and entities involved in the study.

2. Identifying the most appropriate way the data can be clustered into, tagged and elaborated to respond to the purpose of the study. The other factor that increased the variability of the rail unit cost in previous studies was the lack of uniformity within the sample of projects analysed in terms of characteristics of the investments. Thus, the project information has been sorted in a way to tackle the scarce uniformity of rail infrastructure projects. This was achieved by - on the one hand- defining clusters of projects sharing similar characteristics – and on the other hand- breaking down investment cost and analysing similar cost categories (e.g. referring to the same components).
   To accurately define clusters, information on projects was collected in detail both concerning cost as well as the technical characteristics and the external conditions that can impact on the project cost. The organisation process and the final data set structure are outlined in details in the section 4.1.
3. Defining the way the data are retrieved and aggregated into statistics and, coherently, developing the interface for such information to be easily accessed, presented and for the database to be populated over time. Different statistical approaches were tested, to identify the most appropriate and reliable one. This was used as the numerical basis for the development of the REGIO Rail Unit Cost Tool, as described in section 5.1.

The following sections present the way data was collected, cleaned, manipulated, and aggregated and, following, how the REGIO Rail Unit Cost Tool was designed and implemented.

4 Data background

4.1 The data definition process

The first step of the database creation entails the *data definition process*, or more specifically the definition of the data classification method in terms of project type, investment cost and technical data, which are necessary for the tool to perform the project analyses.

The data to be collected was defined considering two opposite needs. On one hand, this refers to the need to collect the highest level of detail possible on projects and, on the other hand, to the need to ensure that the data collection is not burdensome for the stakeholders. The information to collect regarding technical characteristics and the different cost items included in the investments were carefully assessed together with experts and the Commission, to ensure the right balance. Data collection was thus organised following a progressive multi-layered iterative process.

As concerns the technical characteristics, a number of features that impact on rail infrastructure project unit cost were identified from the literature available and with the support of railway experts. Technical characteristics were organised considering that they have different levels of impact on the investment cost.

The first and primary clustering dimension was identified in the type of works executed. Four sub-clusters consisting in different intervention categories were identified together with experts, stakeholders as well as the Commission and are outlined in the table below.

Table 5 - Clustering of observations on the basis of the type of works inlcuded in the investment

Project category	Description
New lines	Construction works related to the deployment of a new railway line.
Rehabilitation	Construction works mainly related to improvement work that are conducted to reinstall the design characteristics of an existing line (e.g. design speed, capacity, etc.) which has worsened in quality due to a lack of sufficient maintenance. Rehabilitation can also refer to works that are executed to make an existing line compliant with new EU regulations.
Upgrade	Construction works related to improving the original design parameters of an existing line, usually in terms of capacity or speed.
Signalling, telecommunication and electrification	Construction works related to the deployment of a signalling, telecommunication and/or electrification system along a railway line.

While the work type was identified as the main source of differentiation across projects, railways operated by high speed trains and those operated by conventional trains significantly differ in terms of **technical specifications**. Such technical specifications have a significant impact on the costs of each work type, which must comply with different standards and safety requirements[19] (high speed infrastructure costs more than conventional railways).

Specifically, railways have been classified according to EU Council Directive 96/48/EC as:

[19] See Decision 2012/462/EU

- **high-speed** is classified as newly built infrastructure that can be operated at a speed that is equal to or higher than 250 km/h or that results from an upgrade of a pre-existing line which can then be operated for speeds of at least 200 km/h;
- **conventional** is classified as newly built infrastructure that can be operated at a speed lower than 250 km/h or an existing line that can be operated at a lower speed than 200 km/h. [20]

A list of the other cost-impacting features has been developed and validated with experts. Accordingly, information on the technical characteristics of rail infrastructure projects has been organised, according to progressive level of details (see Figure 4).
A detailed analysis of the impact of each feature is reported in Annex 9.

The same approach is applied to the investment cost data assessment (**estimated** and **final**). Three levels of analysis were identified, corresponding to a progressive breakdown of cost, which is divided in sub-components. The first level concerns the analysis of the general project cost elements (i.e. costs relative to planning and design activities, ancillary costs and construction costs) and progresses into higher details at the second level until reaching the analysis of each single cost item at the third level of analysis.

The information on technical characteristics and cost collected for each level of detail are presented below.

4.1.1.1 BREAKDOWN OF INVESTMENTS INTO SUB-CATEGORIES OF WORKS

The organisation of the information was designed through a **three-tier approach**. Each tier corresponds to progressive level of detail of the information on project costs and technical characteristics:

- the **first tier** includes general information on the technical characteristics of the projects (e.g. line category, geographical scope of the investment, etc.) and a preliminary categorisation of the total project costs into three sub-categories (planning & design, construction works and ancillary costs);
- the **second tier** focuses on two main parts of the rail infrastructure project - construction works (and associated costs) and ancillary costs.
- the **third tier** focuses on costs and technical characteristics relative to the equipment and permanent way, and tunnels, which are classified based on their construction method.

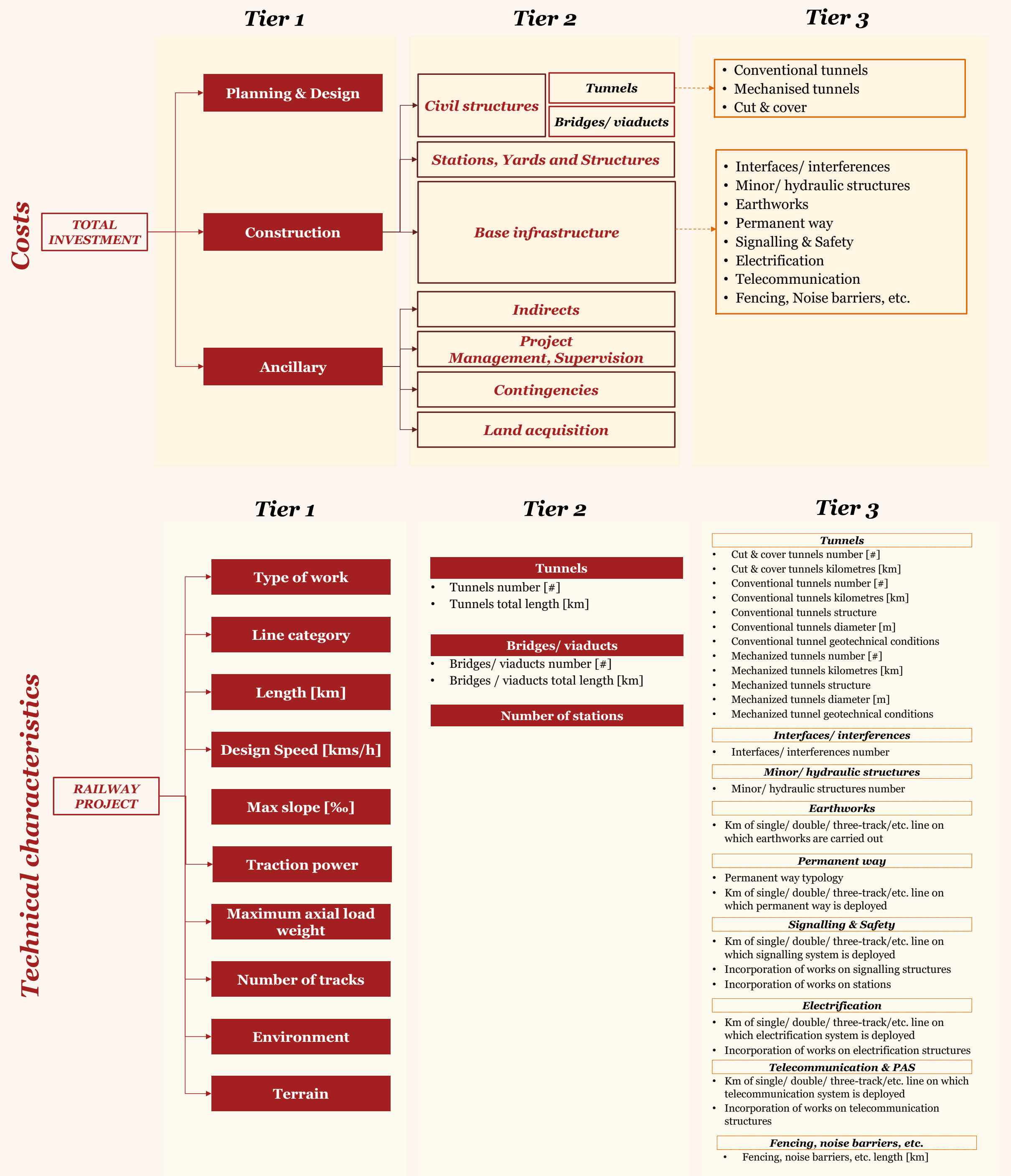
This approach allows carrying out a series of analyses, increasing their accuracy progressively by increasing the level of detail, focusing the analysis on a specific cost element.

It has to be noted that the classification of the information in the three tiers follows a logic of increasing level of detail and does not necessarily correspond to an increased detail of the cost estimate (e.g. information on the tunnels located along the railway line (which is included in tier three) may be elaborated during the preliminary planning phase).

Figure 11 presents the structure of the three-tier approach, for both cost and technical specifications.

[20] Under the definition of Directive 2004/50/EC of the European Parliament and of the Council of 29 April 2004.

Figure 11 - Tier Structure



The exact definitions of the identified cost categories, which were considered in the compilation of the database, are presented in Annex 3.

The database was designed following the approach explained above in order to collect and organize all the information gathered in the process in a sensible way and coherently with the scope of the study.

4.2 Data collection

The study is strongly dependant on the level of information collected in the database. Coherently, a well-established methodology for data gathering, organisation, cleaning and manipulation was designed. It entails the use of existing databases, the involvement of stakeholders as well as the review of information using third-level sources.

Different sources and approaches to gather data have been combined, to check, integrate and complement one another. Below, the data collection sources are outlined.

4.2.1 DG REGIO and INEA databases

The data collection process started from gathering the quantitative information from the data sets made available by DG REGIO and INEA. Specifically:

- DG REGIO provided technical documentation, project applications and Final Implementation Reports related to 365 projects supported under three programming periods (2000-2006, 2007-2013 and 2014-2020) as well as the DG REGIO database (Work Package 10), which, however, could not be used for the analysis due to the limited reliability of the data included[21].
- INEA made available:
    - an electronic database including information and costs for the activities envisaged in each project supported under the TEN-T Programme;
    - an electronic database including data on projects supported under the CEF Programme over the years 2014-2015, as well as the project applications and the related grant agreements.

The extraction of the information consisted in a desk analysis of all the documents available for each project included in the analysis. Thus, it results in a data mining and disaggregation exercise performed on a project-by-project basis.

The Data Mining Process in Numbers

More than 500 projects have been analysed in detail. Depending on the documentation provided (mostly in form of project application), each of them comprised of approximately ten to 30 documents to be investigated with up to 300 pages for individual documents. This activity involved ten PwC experts who has been assigned to this task over a three months period.

As information have been collected in different languages (based on the language of the project application), PwC teams from different locations have been involved.

Due to the size of the data to be processed, each expert spent, on average, **five hours per project** to compile all relevant information from the project application documents.

In total, over 300 person-days have been used to complete this activity.

4.2.2 The stakeholder consultation

Stakeholders have been involved in a thorough consultation process to integrate, revise and complement the data collected from DG REGIO and INEA databases. Specifically, the involvement of the stakeholders

[21] The cross-checking of the information included in the database have shown that, out of five projects analysed, the data was incorrect for three projects.

aimed to both integrate information on projects for the quantitative analysis on unit costs as well as to identify country level specificities and factors impacting on the cost of railway construction.

The stakeholder consultation has been carried out providing full support to stakeholders in the collection, organisation and filling-in of data, activating local experts to ease the burden for stakeholders and to ensure the standardisation of the process throughout Europe. The details of the consultation process are provided in Annex 1 and Annex 2.

STAKEHOLDER CONSULTATION

The stakeholder consultation has been designed to collect:

- qualitative information on country level specificities and factors that may impact project costs (e.g. market conditions, legislation, procurement practices, cost estimation methodologies, etc.);
- quantitative information on the projects under analysis, regarding technical characteristics and costs (estimated and final).

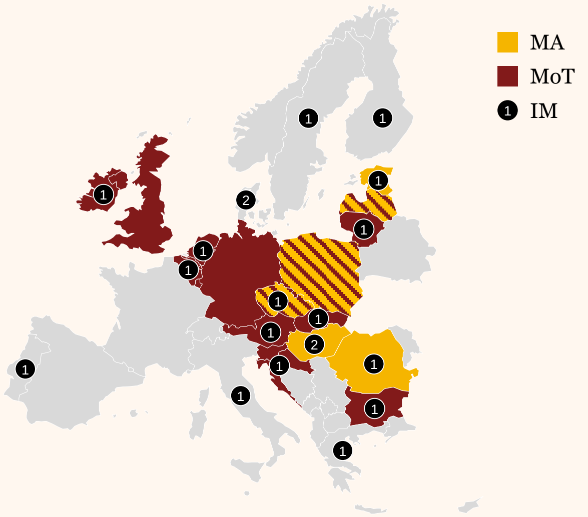
While the former has been gathered by consulting different stakeholders involved in the design and implementation of rail projects (e.g. Ministries, Transport Agencies, IM, MA, etc.); the latter has been primarily provided by IM.

Results of the qualitative consultation

The qualitative consultation was distributed to 124 organizations throughout Europe and lasted approximately 5 months.

Overall, 41 different organizations took part in the consultation: 20 IM, 6 MA and 15 MoT.

Figure 12 – Participation in the qualitative consultation



Results of the quantitative consultation

The quantitative consultation involved 35 IM, which have been in charge of the implementation of rail infrastructure projects over the years under analysis.

These organisations have been asked to provide information on the technical characteristics, estimated costs and final costs of the projects carried out.

Considering the relevant amount of information required, many IM have been supported by PwC local teams in retrieving the required information, as outlined above. Additionally, to reduce the effort required, in case of a high number of projects, priority has been given to those carried out in the 2007-2013 programming period and the level of detail has been reduced (e.g. by asking information up to Tier 2, excluding Tier 3).

The consultation enabled the collection of detailed information on 94 projects.

The detailed information on the activities related to the consultation of stakeholders are provided in Annex 1.

The quantitative stakeholder consultation: a communal European effort

Stakeholders have been thoroughly involved in the data collection phase. The data necessary for the development of the database were indeed held by local organisations and authorities. Most collaborated actively in the progress of the study.

For six countries (the Czech Republic, Greece, Latvia, Lithuania, Poland and Slovakia), local PwC teams have provided hands-on support retrieving and extracting data from archives. PwC local experts supported the IM both collecting the information as well as translating this data and entering it into the database in the appropriate format. Such activity was carried out to support the IM, which otherwise would not have been able to complete the workload in time for the study to use the data.

Furthermore, telematics support was provided, to clarify any doubts stakeholders encountered. In total this activity took around **175 person days** to complete, including the extraction of the data from the stakeholders and the initial consultation, as well as the reminders and support with filling out the questionnaires.

Annex 2 presents the list of stakeholders that participated in the data collection exercise. The Annex also reports the extensive work done to identify the appropriate experts within stakeholders' organisations and liaise with them to gather the information requested.

4.2.3 Literature review and desk analysis

In parallel with the data collection process using inputs from the stakeholders, information has been collected from literature as well as through desk analysis. In particular:

- the information previously collected has been checked against documentation available on the internet,[22] as well as using the documentation available within PwC internal databases.[23] In case of inconsistencies within data provided by stakeholders, they were asked to either provide clarifications or revise the data.

  Concerning the documents included in the databases made available by DG REGIO and INEA analyses were carried forward to retrieve information on project technical characteristics and costs to be included in the dataset. A focus desk research has been carried out on out-of-range observations, to check the reliability of the information available in the dataset and to search for additional details on project costs and characteristics.

- Besides supporting the refinement of the methodological approach of the study, the literature review has been used to retrieve information to complement the study dataset. For example, some of the information available from the studies and other literature (such as information from PwC's previous projects on rail infrastructure) has been used to complement elements of the data base, as it provided information on further rail infrastructure projects.

The bibliography in annex reports the main studies and literature used in the progress of the study.

[22] E.g. Press reviews, associations and organisations of railway stakeholders, project-specific websites, etc.

[23] Information are mostly relative to previous studies and assignments, which have been carried forward by PwC.

4.3 Data cleaning

The data gathered and reviewed was cleaned of all those observations that could not fit the scope and purpose of the analysis. This particularly refers to:

- Observations entirely related to structures that were beyond the scope of the study (e.g. investments related to the construction or upgrade of railway stations, construction of inner-urban railway lines, investments in rolling stock, studies, etc.).
- Observations from the databases provided by DG REGIO and INEA where it was not possible to retrieve at least the minimum level of information for the analyses (i.e. total cost of the project, length of the infrastructure and basic information on the work type).
- Observations for which clear inconsistencies were encountered in the data breakdown and for which it was not possible to perform a data review.

Out of the 867 projects included in the DG REGIO (365) and INEA database (502), approximately 55% concerned rail infrastructure projects that were within the scope of the study (see section 1.2). This database was complemented by additional projects provided by stakeholders and projects from PwC internal sources. Another 10% of the projects (80) was then dropped due to insufficient information or misleading data. This resulted in a database of approximately 400 projects. Nonetheless, many of the EU funded projects related to different phases of the same project, wherefore those were combined to single projects with complete information on the entire line sections. The whole process, which is described in details in Annex 8, led to the creation of a final database embracing data on cost and technical characteristics of **208 rail infrastructure projects** which is outlined in section 4.4

4.3.1 Data Normalization

The final database includes projects from different countries and from different time periods. Such information has been therefore normalized in order to perform the analysis on comparable project data.

Based on the literature review carried out and the analysis of the methodology followed by similar assignments, two possible approaches have been identified and used for data normalisation:

- **Normalisation based on Purchasing Power Parity (PPP) rates,** aims to convert all costs in a fictional currency that represents the (weighted) average of the EU28 countries in terms of purchasing power parity (EUR EU_{28}).[24]
- **Normalisation based on inflation,** aims to adjust costs incurred in different years to values relative to one base year. The inflation adjustment factors are based on the producer price index of each country.

The application of both approaches presents pros and cons, which the study attempts to overcome.

Specifically, PPP exchange rates are widely used also in other rail construction benchmark studies,[25] as they represent a complete and thorough indicator to account for differences in the value of money across different countries (also countries using the same currency). The rates are developed based on the amount

[24] For MS where a currency different form EURO is used, costs are preliminarily converted from the local currencies to US$, based on the PPP rate made available from OECD. Values in US$ are then converted in EUR based on the conversion ratio between US$ and EUR.

[25] Purchasing power parity rates are often used in literature when comparing construction costs across economies. Examples include: Kenneth D. Walsh, M.ASCE; Anil Sawhney, A.M.ASCE; and Audrick Brown IV, S.M.ASCE 2005 International Comparison of Cost for the Construction Sector: Purchasing Power Parity, Journal of Construction Engineering and Management, vol. 131 issue 2, Feb.

Rick Best, Jim Meikle 2015 Measuring Construction: Prices, Output and Productivity
PwC 2016 HS2 Phase Two capital costs against other international high speed rail projects.

of money that consumers require to purchase a basket of goods and services. This basket includes goods (and services) consumed by individuals and private households, and therefore may not be entirely accurate to account for differences in rail construction cost among countries. In addition, the PPP exchange rate is less effective when prices tend to vary strongly across years rather than across countries.

On the other hand, the normalisation based on inflation prices takes into consideration the evolution of the prices over time. The producer price index measures price changes from the perspective of the seller, using the average change in selling prices indicated by domestic producers over time. This index estimates the production costs and their differences over different years. The reliability of the normalisation based on an inflation index is higher when the evolution of only one economy is analysed. Indeed, this approach does not enable for consideration of the differences among countries, but is only able to provide for normalisation of changes over time in the same country. Since this analysis on rail unit cost focuses on a variety of countries with different inflation rates, the results would lead to the underestimation of cross-national differences.

In order to respond to the needs of ensuring a normalisation both across regions as well as throughout time, the normalisation has been carried out following a combined PPP-inflation approach. A mixed indicator was therefore developed to consider both the differences among countries and the impact of time on the infrastructure cost. The combination of these factors was carried out taking into account the proportion of railway infrastructure costs as provided by experts in railway construction: generalising to the extent possible, country specific labour cost makes up approximately 20% of the total cost, while the remaining 80% of the cost refer to goods or services which can be bought on a shared international market (e.g. raw materials).

The labour cost was normalised using an index based on PPP. This is based on the assumption that salaries paid in different countries for the same profession are able to purchase the same amount and quality of goods and services provided the different price levels of the different countries. The remaining 80% of the cost has been normalised using the producer price index. Since suppliers for rail materials are usually large companies operating on an international level and these goods and services are available on a shared European market, it is assumed that there are no differences in price between countries, but only among different years.

The creation of the normalisation factor

The PPP exchange rates made available by the OECD are based on the comparison of each currency with the US dollar of each specific year of the reference period, whereas the index needed for the study should be based on Euro and should be normalised to account not only for differences across countries, but also across time. Therefore, starting from the OECD index, a normalisation factor was derived:

1. For MS where a currency different from EURO is used, costs are preliminarily converted from the local currencies to US$, based on the PPP ratio made available from OECD. Values in US$ are then converted in EUR based on the conversion ratio between US$ and EUR.
2. The indicator of the €PPP has been multiplied by the European Harmonised Index of Consumer Prices (HICP) to consider the differences of over different years.

4.3.2 Harmonisation of technical data

To ensure that investments could be compared using the same unit (e.g. EUR / km of line), it was necessary to account for the differences in line characteristics (i.e. number of tracks). A number of projects refers in fact to both single and double track railway lines, others to only double-track and a limited amount to only single-track.

A standardised length, which translates the railway length into a double-track equivalent, was therefore used to make the data comparable. All the analyses refer to this measure when considering the size of the infrastructure.

4.4 The database

For matters of simplicity and universality of use, and to comply with the requests of the Commission[26], the database which constitutes the basis for the REGIO Rail Unit Cost Tool was designed using MS Office Excel. Following the results of the data collection activities, it was populated by **208 rail infrastructure projects carried out in Europe over the 2000-2015 period.**

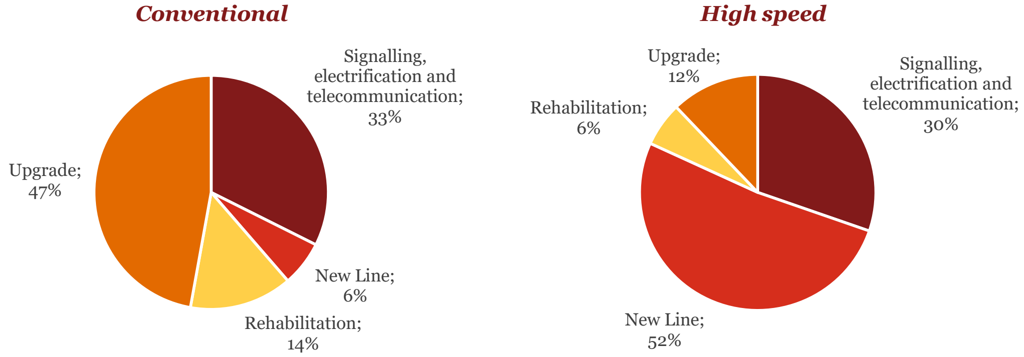
The data included in the database represents a significant sample of the European rail infrastructure market. Indeed, the projects analysed in the study account for approximately **50,000 kilometres of infrastructure investments** (of which ca. 24,000 km were included in the final version of the database),[27] which represent works on **over 20%** (10% in the final database) **of the entire EU railway system.**[28] Additionally, a complete geographical coverage is ensured, as data on the investments has been collected for projects from **all 26 EU Member States with rail infrastructure,** with the only exception of Ireland.

The main characteristics of the database are further outlined in the following paragraphs.

4.4.1 Type of works and line categories

The majority of the projects included in the database (175 out of 208) relate to works performed on conventional railway lines. Only 33 projects refer to works carried out on high-speed lines. This difference is related to the different distribution of infrastructure categories for each type of line. While conventional lines works are in fact generally related to rehabilitation, reconstruction and upgrade interventions and therefore carried out on smaller sections, high speed line projects may encompass hundreds of kilometres of new railway line.

Figure 13 – Projects in the database by work type and line category



This distribution reflects the trends in rail infrastructure investments outlined in section 2. The current disproportion in the distribution of projects along the four work types is expected to be flattened in the future, given the possibility to expand the database (as outlined in section 5.1.3).

A SIGNIFICANT SAMPLE OF THE RAIL INFRASTRUCTURE INVESTMENTS IN THE EU

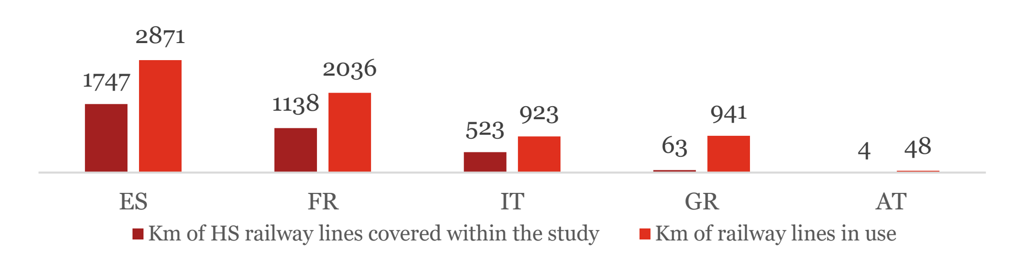
With regards to the high speed project, the analysed investments represents almost **45% of the high-speed lines currently in use in the EU**. The investments under analysis are mainly located in Spain, France and Italy, covering approximately 60% of the high-speed railway lines in use in these countries (see figure below).

[26] As outlined in the Terms of Reference attached to the contract 2013CE16BAT064

[27] Including new lines being constructed, upgrades, electrification works, signalling works, etc.

[28] Eurostat data report approximately 215,000 kilometres of railway lines.

Figure 14 - Coverage of the study regarding new high-speed lines

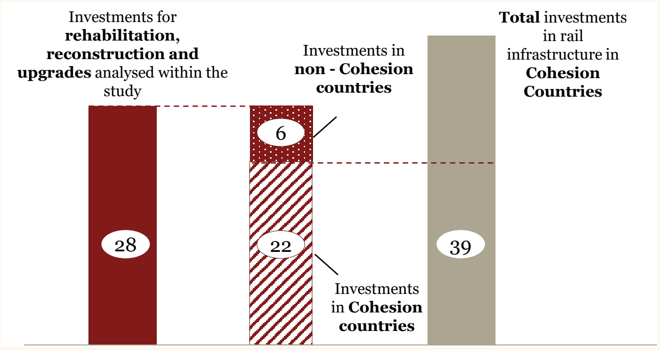


Source: PwC elaboration on data from the Study database, OECD and the Rail Market Monitoring of the EC[29]

The assessment of the single investments highlighted that almost 55% of the projects refer to rehabilitation and upgrade works. This phenomenon can be explained as upgrades and rehabilitations usually relate to rather small sections of lines (ca. 55 km on average). Nonetheless, the study covers upgrade and rehabilitation interventions carried out on approximately 6,300 km of railway lines.

As stated in section 2, the majority of renewal and upgrade interventions have been carried out in Cohesion Member States between 2000 and 2015. Coherently, the majority of investments for rehabilitation and upgrade considered in the analysis (ca. 79%) are located in these countries, and correspond to almost 60% of the total rail investments undertaken in the same territories between 2000 and 2015.

Figure 15 – Coverage of the study regarding rehabilitation and upgrades (data in EUR bn)



Source: PwC elaboration on data from the study database and OECD

As introduced in section 2, the European Union has defined unique standards to be applied in all Member States in terms of signalling and train control notably: the ERTMS. The regulatory requirement to comply with the interoperable, European system led to significant investments being made to ensure the compatibility of existing lines – as well as newly constructed ones – with the standards. As a result, investments in signalling have been significant in the railway sector, and, to our scope, account for approximately EUR 4.6 billion, out of which ca. **EUR 1.1 billion** refer to **deployment or upgrade** of **ERTMS systems.**

4.4.2 Level of detail of the information on investment cost

Data on the investment cost of each project has been collected and organised in order to reflect the three-tier approach (see previous section).

[29] Fifth report on monitoring development of the rail market, EC, 2016

Information on total **estimated investment costs** and total lengths have been collected for all the projects in the database. For 90% of them, the total investment cost was additionally split between cost related to planning and design, construction cost and ancillary cost and general information on technical specifications. Whereas, information on technical characteristics and estimated costs at Tier 2 and Tier 3 level was collected for 128 and 114 projects respectively.

Information on **final cost** has been collected for approximately 40% (78) of these projects, also because part of the investments included in the analysis are still under construction. For the large majority of them (65 projects), it was also possible to gather data on the final cost of construction, planning and design and ancillary cost. Whereas, detailed information was collected only for approximately 50% of them (i.e. information on final cost at Tier 2 was collected for 44 projects, at Tier 3 for 41 projects). Of the 78 projects with total final cost data available, five outliers have been excluded from the analysis, as those do not represent statistically significant cases.[30]

The analysis of final cost

Retrieving information on the final cost of the single interventions included in a rail infrastructure project presents some difficulties as the contracts for the project implementation usually refer to sections rather than elements of a railway lines. Additionally, the information on the final cost of a project may be considered as a sensitive information which IM is not available to share.

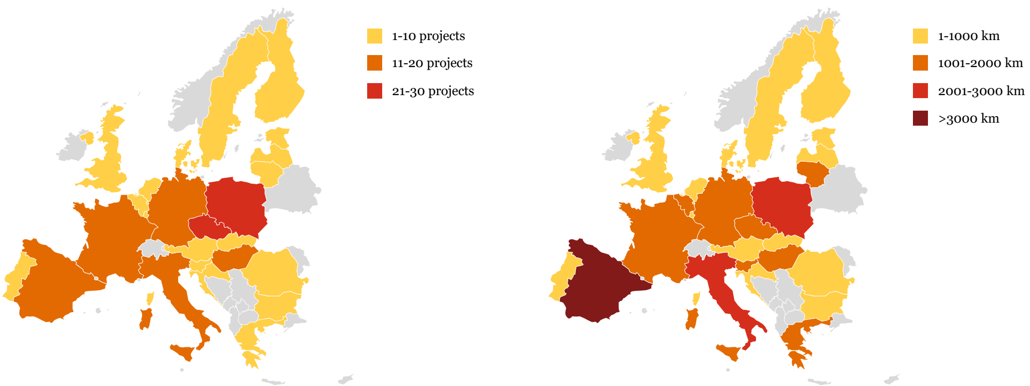
Considering the amount of data available currently in the database, the tool gives to the user the possibility to insert information on **final cost of a projects already present in the database and/or to provide the final cost for a new project.** Hence, users are able to decide if he or she would to carry on a detailed analysis of the investment cost only on expected cost or on both, estimated and final. In addition, the data on the final total cost of the projects have been considered to derive indications on the expected variations of planned cost (see section 5.1.2).

4.4.3 Geographical coverage

In line with the scope of the study focusing on the EU, data has been collected for projects from **all 26 EU Member States with rail infrastructure**, with the only exception of Ireland (for which limited information was collected and then discarded in the data cleaning process described above).

The projects concerning rail investments used in the analyses[31] are distributed as follows (see figure below).

Figure 16 - Projects by country & length of the railway network analysed per country



[30] These projects focused on the implementation of new technologies for which no prior experience and expertise was available, thus resulting in a cost overrun of more than 100% (three of those were the first Italian high-speed rail projects and one each was a Dutch and a Belgian ERTMS implementation project).

[31] Projects related to stations, hubs and bypasses, urban lines, phases and specific civil structures are not included in the analysis of unit cost.

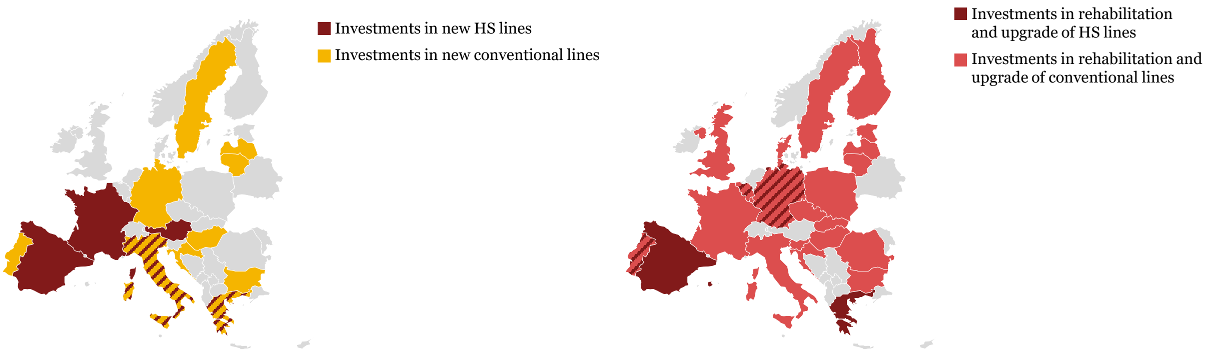
Source: PwC elaboration on data from the study database

Poland is represented by the largest number of observations in the database (28), corresponding to 13% of the total, followed by the Czech Republic with 27 projects and Spain with 15 projects. Also Germany, Hungary, and Italy were strongly represented with 13 projects each. Together these six countries make up half of all observations in the database (about 52%).

Considering the total kilometres of railway lines analysed at country level, the situation slightly differs; Spain is the country represented by the largest amount of kilometres of railway included in the database with around 3,350 km (14% of the total DB), followed by Poland with around 3,000 km (13%) and Italy with approximately 2,200 km (9%). Also Germany, France and Lithuania are strongly represented with circa 1,900 km (8%), 1,800 km (8%), and 17,000 km (7%) respectively.

From a technical perspective, the geographical distribution of the investments analysed in the study reflects these differences between the European Countries (see figure below). A few countries concentrated investments in constructing high-speed lines: e.g. Italy, Spain, and France. Others have mostly invested in renewing and upgrading existing lines, in particular Eastern European countries. Others again, present a mix of new lines and upgrading/renewal of the existing ones.

Figure 17 - Geographical distribution of the infrastructure investment types



Source: PwC elaboration on data from the study database

5 Developing the REGIO Rail Unit Cost Tool

The aim of the REGIO Rail Unit Cost Tool is to aid the assessment of investment costs in rail infrastructure projects, and supporting the decision making process regarding their approval and financing. Therefore, it serves as a feasibility and options analysis tool to evaluate different projects and their funding applications.

Specifically, the tool is able to perform a benchmarking analysis and verify whether the investment cost of a rail infrastructure is comparable to the investment cost of similar projects. These are identified on the basis of the factors that impact on the investment cost as the **technical characteristics** of the railway line (e.g. category of works, line category, structures included in the investment, etc.) and the **external conditions of the project** (e.g. geographical features of the construction site, presence of urban environment, etc.).

Therefore, the REGIO Rail Unit Cost Tool allows for:

- calculating the unit cost of the different cost items which build up the investment cost (e.g. base infrastructure, signalling, permanent way, etc.);
- presenting the positioning of the overall infrastructure in comparison with average assessed projects;
- providing a set of indicators, depending on the level of breakdown of the inputs, per infrastructure category type identifying eventual criticalities if single elements appear out of average ranges.

Strategic planning activities, project option analysis and examination of project proposals can be supported by the use of the REGIO Rail Unit Cost Tool. Hence, the RRUCT in addition to the benchmarking activity can provide a forecast of a project costs. The forecast will base his analysis only on technical characteristics of the railway line that the user will provide. The estimated project cost reported, following the methodology applied in the benchmark, refers to the median value of the cost of projects with similar technical characteristic that are already included in the database.

It is worth highlighting that the tool is designed as a **support** for **benchmarking and assessment activities** and **cannot replace the cost analysis of the construction works which shall be carried out during planning and design phases.** Nevertheless, the RRUCT cannot substitute a detail assessment of an investment and it does not include a risk assessment exercise.

The REGIO Rail Unit Cost Tool is <u>not designed to</u> identify *wrong* cost estimates in application forms for European Commission financial support. The tool is designed to spot unexpected cost variations in rail infrastructure projects, which are not included in the usual range of cost variations of similar projects. It does not mean that a project with unexpected high or low unit cost values is poorly evaluated; it may nonetheless require detailed explanations to make sure that the reasons leading to such exceptional unit cost values are known and can be duly controlled.

The tool has several opportunities for further development, which have been outlined in the following chapters.

5.1 Methodology

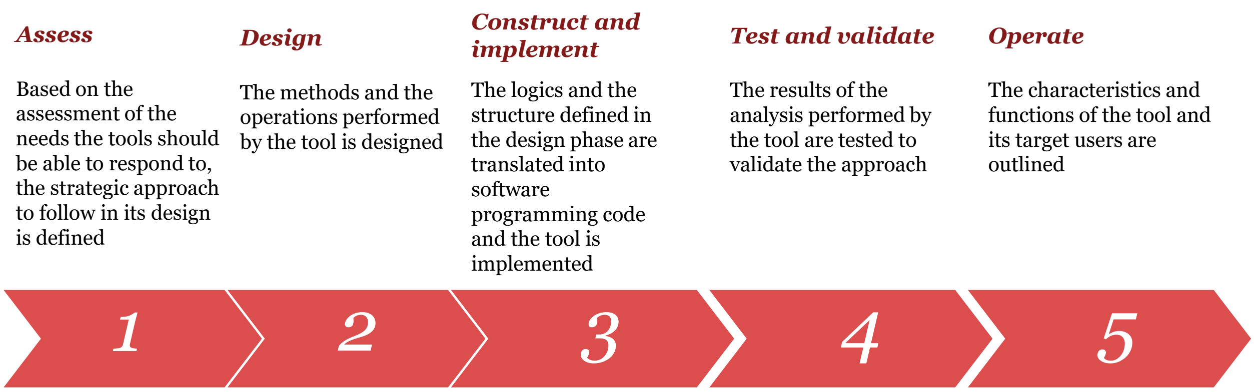
The introduction of the REGIO Rail Unit Cost Tool will change the process followed to assess the cost of rail infrastructure investments. Therefore, the methodology adopted for its development and implementation is the worldwide applied PwC Transform Framework, the methodological framework that encompasses all steps related to a change initiative, from the strategic to the operational stages.

Following a global methodological framework reinforces the methodological approach used for the development of the tool. Moreover, it ensures that all the relevant activities and aspects involved in this process are properly assessed and taken into consideration.

The activities performed to develop the REGIO Rail Unit Cost Tool are grouped to five main stages of the PwC Transform Framework:

1. **Assess.** This stage consists in the assessment of the needs to be met by the tool and the definition of the strategic approach to be followed in the design phase.
2. **Design.** This stage includes the definition of the methods and the operations to be performed by the tool.
3. **Construct & implement.** The main purpose of this stage if to successfully implement the tool designed.
4. **Test and Validate.** In this stage, the results of the analysis are validated by testing the approach and the REGIO Rail Unit Cost Tool.
5. **Operate.** In this stage the way the tool will be operated and maintained once delivered is defined.

Figure 18 - The stages of the PwC Transform Framework



A detailed description of the activities carried out at each stage are is provided in the following paragraphs.

5.1.1 Assess

The purpose of this stage is to identify the strategic approach to follow in the creation of the REGIO Rail Unit Cost Tool, given:

- the objectives the tools is designed to achieve;
- the data background of the analysis, which consists of the database described in section 4.4.

Different approaches have been tested in previous attempts to analyse the unit cost of rail infrastructure projects (see 3.1). This study partially relied on the experience from the past, to test how different methods could produce results that are valid, reliable and useful for the creation of the REGIO Rail Unit Cost Tool.

Based on the review of the relevant literature, two opposite statistical approaches applied to rail infrastructure costs have been identified. One approach is based on the formulation of a **universal law** to be applied to all observations. The other is based on a **bottom-up approach** consisting of an analysis of the different sub-components of the investment cost.

Regardless of the tools used, the first case refers to the attempt to generate a statistical distribution which can capture the majority of the cost values in the sample and then be inferred as a model for calculating costs in any theoretical specific project population, and which thus can be equally applied to all different work types. The approach is generally pursued by applying techniques designed to analyse big data panels. The benefits of this approach are relative to the usability of the outcomes as well as to the possibility to treat enormous amounts of information with a general, unique approach. However, this approach is very sensitive to the amount of data that are used to fill-in the system, as well as to the quality and level of detail of the information included in each single observation. **Therefore, the universal law is not ideal for the development of the REGIO Rail Unit Cost Tool as it can be hardly applied to small data samples.** Nonetheless, the possibility to apply a methodology following the big-data approach has been investigated, by testing the possibility of applying the **attribute selection process** and of **running a regression analysis**.

The attribute selection process

The **attribute selection process** consists of the use of a software to search automatically for a set of relevant features that impact the unit costs and should be used in the model construction. To this aim, it represents the very end of the spectrum of the analyses undertaken, as no human intervention was made to indicate which type of relation was to be identified within variables. Indeed, the approach is based on reiterating correlation analyses without accounting for the logical expectation of relationships among variables.

The process has been performed on the whole dataset including not only the information on project technical characteristics, but also indicators related to the countries where the projects were deployed:

- the score in the Corruption Perception Index, developed by Transparency International;
- the score in the Competitive Industrial Performance Index (CIP), made available by the United Nations Industrial Development Organization;
- the rate of urbanisation, as reported in the Urbanisation index of the United Nations Population Division;
- the share of mountain territory, retrieved from the report "Mountain Areas in Europe: Analysis of mountain areas in EU member states, acceding and other European countries" (Nordregio, 2004);
- the number of inhabitants, retrieved from Eurostat;
- the total kilometres of rail infrastructure in the country, as recorded by the World Bank;
- the number of annual railway passengers, based on Eurostat data;
- the number of train kilometres, based on Eurostat;
- the density of the population in the project area, based on data from the TENtec system.

However, the results of the analysis - performed using the *Weka Software* - did not shown any significant outcome. Specifically, testing the CorrelationAttributeEval algorithm, which evaluates the validity of an attribute by measuring the correlation between one attribute and the remaining set of attributes through the Pearson distance metric, identified a high number of attributes (12 attributes) with a limited correlation (ca. 4%) with the unit costs. Thus it was not possible to identify a relevant set of attributes to be considered to deduct the unit costs.

Regression analysis

Additionally, the possibility to run a **regression analysis** has been investigated. Both OLS and weighted regression techniques have been used to isolate the statistical relationships between one variable (the cost of the infrastructure) with multiple factors that identify the complexities relative to the infrastructure type, orography, etc. The analyses have been performed on the observations classified based on the infrastructure categories and based on the line category (i.e. conventional and high-speed lines). For the purpose of our analysis, the central hypothesis is

$$\mathbf{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \qquad \textit{(Equation 1)}$$

Where $\mathbf{y}$ is the vector of infrastructure cost, $\boldsymbol{X}$ is the design matrix, which represents the elements that impact the cost of the infrastructure, $\boldsymbol{\beta}$ is the vector of the coefficient, which determines the degree of impact of the regressors to the dependent variable and $\boldsymbol{\varepsilon}$ represents the error term.

The main limitation to the regression analyses performed was identified in the quantity and quality of data at disposal. Indeed, the regression function requires that for each observation there is non-null information relative to all the elements included in the function, otherwise the entire observation cannot be used and is dropped. This issue significantly limits the level and the degree of analyses that can be performed, reducing either the number of observations or limiting the number of variables that can be used to identify the correlation among the dependent variable and the independent ones.

Further, the analysis of the distribution of the residual values highlighted the presence of very elevated **heteroskedasticity**, which compromises the validity of the results.

The approach based on the identification a unique, general law for all rail investments did not produce reliable results that can be used for the design of the REGIO Rail Unit Cost Tool.

A bottom-up approach, which aims to determine the cost of different cluster increasing progressively the level of detail, was subsequently investigated. The approach is not entirely new, but recalls those that have been applied to some of the most fruitful studies on the matter (e.g. DTU and HS2).

The bottom-up approach

Opposite to the previous approaches, which aim to determine a cost value based on statistical relationships with similar projects, the bottom-up approach aims to determine the cost of different cluster of infrastructure works and cost categories.

The bottom-up approach thus consists in identifying the unit costs on smaller portions of the database (either according to specific work types or to specific components of wider infrastructure) and elaborating descriptive statistics on them to derive cost ranges for each rail infrastructure component.

The bottom-up approach has the benefit of producing significant results also with limited data availability and quality, which is the primary limitation encountered in the general, top-down approaches due to the characteristics of the database.

Therefore, the bottom-up approach fits the needs identified in the development of the REGIO Rail Unit Cost Tool. Nonetheless it requires a very thorough analysis being carried out on a project-by-project basis.

5.1.2 Design

This stage defines the logic behind the REGIO Rail Unit Cost Tool and the analysis to be performed.

Firstly, it is considered that the data entered by the user have to be normalised to ensure the reliability of the benchmark[32]. Therefore, the information on project cost and length is normalised following the same approach adopted for the creation of the data background (see 4.3).

As mentioned above, a bottom-up approach is used to build the REGIO Rail Unit Cost Tool and perform the analyses. Such approach is based on the calculation of unit cost of different infrastructure and cost categories. Accordingly:

- the investment cost of rail infrastructure project is breakdown into its subcomponents and the unit cost range is identified for each of them;
- the project are clustered based on the internal and external factors that impact on their cost, in order to calculate the unit cost for uniform clusters of projects.

[32] For the forecasting function, country level differences in the value of goods and services are maintained.

A detailed description of these logics is provided in the paragraphs below "Analysing the components of rail infrastructure investment cost" and "Analysing uniform clusters of projects".

Furthermore, the REGIO Rail Unit Cost Tool is designed to provide an indication of the likely final cost of the investments assessed, as outlined in the paragraph "Predicting final costs".

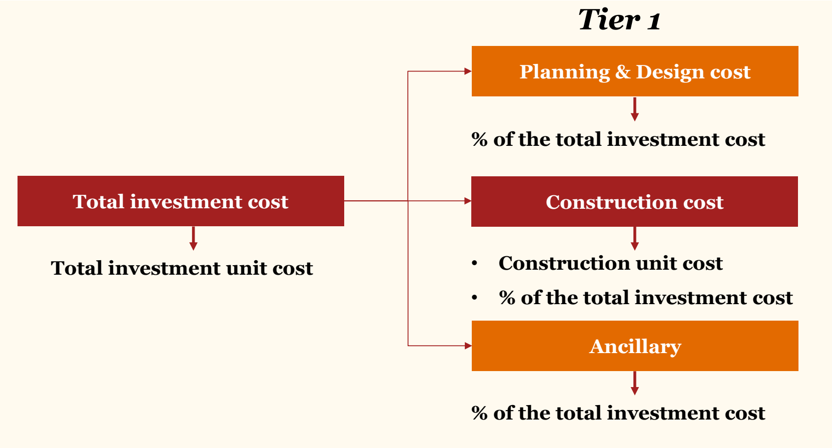
5.1.2.1 ANALYSING THE COMPONENTS OF RAIL INFRASTRUCTURE INVESTMENT COST

The analysis of the investment is designed based on the cost categories identified in the data definition process. Three levels of analysis are defined, following a progressive breakdown of the investment cost.

The first, more general, analysis is performed considering the total investment cost and its main components and comparing them to the project length. Based on this information, the REGIO Rail Unit Cost Tool calculates and analyses:

- the overall unit cost, determined on the basis of the total investment costs and project length;
- the construction unit cost, estimated on the basis of the project construction costs, thus excluding ancillary costs and costs related to planning and design.

Figure 19 – Analysis of the total investment cost and its main components



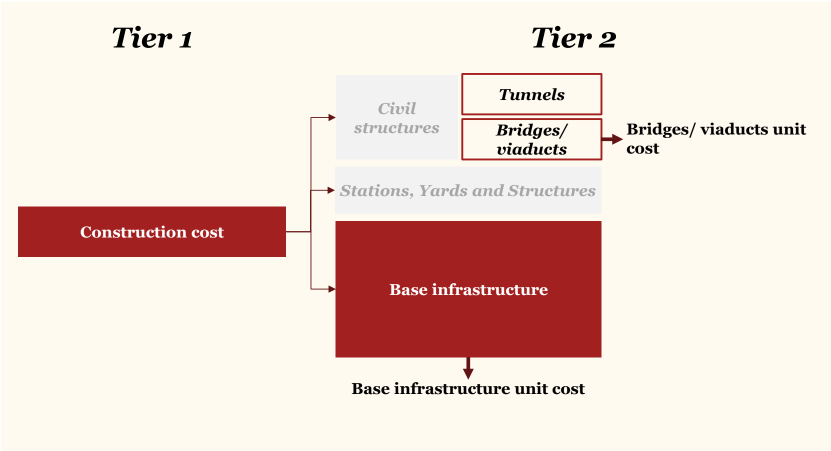
The tool is able to calculate the share of the total investment cost allocated to each components and compare it with the general trends resulting from the database (benchmarking function) or to extrapolate the average costs (forecasting function). This analysis can be particularly significant in benchmarking the ancillary cost. Indeed, in case an unexpected high value of ancillary cost is encountered, this may be related to a high value of the contingencies, thus it is likely that the project encompasses a high level of risk. For this reason, the breakdown of ancillary cost – comprising cost related to contingencies, land acquisition, Project Management and indirect cost - is analysed as well.

The second step of the analysis focuses on the main components of the construction cost. The REGIO Rail Unit Cost Tool calculates the unit cost of the *base infrastructure*, which represents the sum of the infrastructure components relative to the development of a simple, complexity-free railway line. The *base infrastructure* includes construction works and cost elements which represent the minimum set of works to deploy a railway infrastructure. In other words - considering the construction works usually included in a railway project - the base infrastructure is defined as the works (and costs) related to:

- earthworks;
- minor/ hydraulic structures and interfaces;
- equipment and permanent way

The cost of stations[33] and civil structures[34] is excluded from the *base infrastructure* cost, as outlined in the figure below.

Figure 20 – Analysis of the base infrastructure cost



While calculating the unit cost of the stations is not included in the scope, the analysis of structures as bridges/viaducts and tunnels is carried out, but only to provide a general indication of the unit cost ranges for these complex structures. In particular, a dedicated case study was developed for tunnels in order to further assess in detail the technical factors that impact on the costs of these structures (see Annex 13).

The last step of the analysis concerns the single work and cost items encompassed in the investments. Specifically, the REGIO Rail Unit Cost Tool elaborates and compares the unit cost of:

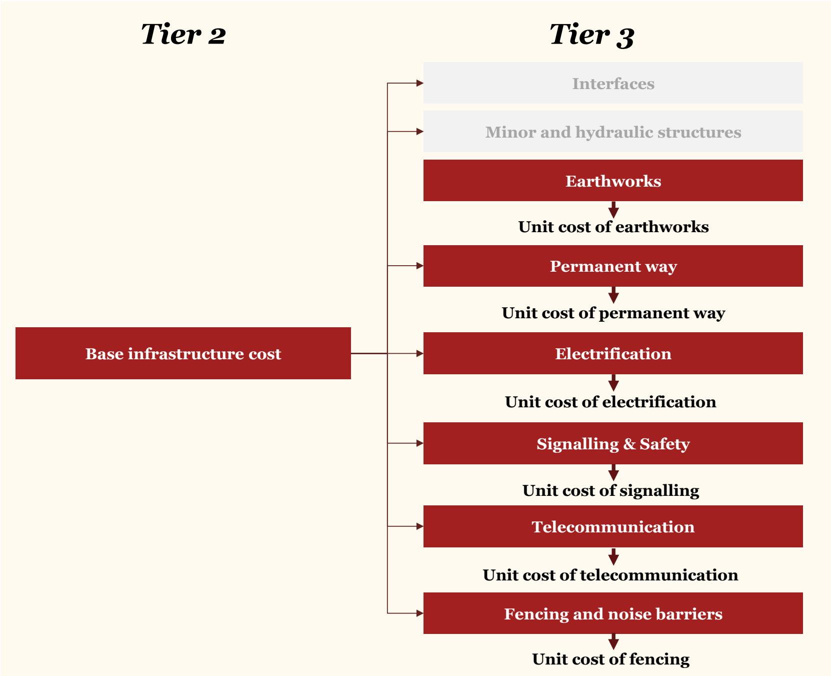
- Earthworks;
- Permanent way;
- Signalling & Safety;
- Electrification;
- Telecommunication;
- Fencing and noise barriers.

In these cases, the unit cost is calculated based on the total kilometres of line on which each category of works is performed.

---

[33] In line with the assumptions considered for the compilation of the database, the cost of stations refer to the cost of related to the infrastructure (Station Yard included), e.g. costs related to the commercial activities are excluded as well as cost of the permanent way and the equipment of the railway line passing through the station.
[34] In line with the assumptions considered for the compilation of the database, the cost of civil structures refer to the cost of construction/ excavation of the structure, therefore, the cost of the permanent way and equipment of the railway line passing through the tunnel/ bridges/ viaduct should be excluded

Figure 21 – Analysis of single cost components



5.1.2.2 ANALYSING UNIFORM CLUSTERS OF PROJECTS

In performing the analysis of unit cost, the REGIO Rail Unit Cost Tool takes into account the impact of two categories of external factors which might cause variation of the unit cost of the different investments:

- Endogenous factors, related to the railway line features (e.g. maximum design speed, category of service, etc.)
- Exogenous factors, related to the sites where the railway lines are deployed (e.g. orography, urbanisation, market conditions, etc.)

Such factors might have an opposite impact on the unit cost (e.g. the maximum design speed and the topography of a railway line may have contrary effects on the base infrastructure unit cost in the case of a high-speed line passing through rural areas). The optimum solution for assessing properly the impact of all the cost drivers is to compare the cost of each investment with that of a sample of projects sharing the same features.

The REGIO Rail Unit Cost Tool implements this solution by translating the factors identified during the data definition process into filters applied to the data background. Their application enables to select the projects of the dataset which match all the characteristics of the investment under assessment (see figure below) or of that for which it is researched to forecast the cost.

Figure 22 – The identification of the projects similar to the project assessed

	Project under assessment	Sample of projects for the benchmarking analysis
Type of works	New line	New line
Infrastructure category	High-speed	High-speed
Design speed	301 km/h – 350 km/h	301 km/h – 350 km/h
Environment	Rural	Rural
Terrain	Flat	Flat
Category of service	Passenger	Passenger
Axial load	> 8,5 ton/m	> 8,5 ton/m

The number of projects included in the database that match all the characteristics of the investment under assessment (or for which it is researched to forecast the cost) is likely to be limited, given the current size of the database (see 4.4). To overcome this limitation, the factors impacting on rail infrastructure cost have been prioritised based on the magnitude of their impact. In case the sample matching all the characteristics of the investment under assessment results to be too limited, the information related to the least relevant factor is discarded in the selection of the clusters. The process is reiterated until a sufficient sample can be identified (see figure below). **A sample that includes more than 40 projects is considered statistically significant and therefore the unit cost range determined is deemed reliable.**

Figure 23 – The creation of the samples

	Project under assessment	Sample 1	Sample 2	Sample 3
Type of works	New line	New line	New line	New line
Infrastructure category	High-speed	High-speed	High-speed	High-speed
Design speed	301 km/h – 350 km/h	301 km/h – 350 km/h	301 km/h – 350 km/h	301 km/h – 350 km/h
Environment	Rural	Rural	Rural	Rural
Terrain	Flat	Flat	Flat	Flat
Category of service	Passenger	Passenger	Passenger	
Axial load	> 8,5 ton/m	> 8,5 ton/m		
		12 projects	**20 projects**	**41 projects**

The details on the prioritisation logic of the different cost drivers for each infrastructure category and cost item is reported in Annex 9.

The box below presents how the clustering approach is applied in the analysis performed by the REGIO Rail Unit Cost Tool.

The creation of uniform clusters of projects for the benchmark and forecasting analyses

For the benchmarking function, the identification of homogeneous clusters, for the assessment of projects, is outlined below, **taking as reference a fictitious project**, which characteristics are listed in table below.

Similarly, also the forecasting function is based on the identification of homogeneous clusters of projects, among those stored into the database. The difference is therefore that, while in the benchmarking analysis one project is assessed against a set of comparable cases, in the forecasting only the projects already in the database are analysed, to understand most likely unit cost values from historical statistics.

Instead of the characteristic of a project under assessment, the case will therefore refer to the characteristic of projects of which we aim to derive an average likely unit cost (overall and at component level)

Table 6 - Features of the project under assessment

Overall characteristics of the project		***Characteristics of the works included in the project***	
Type of works	Upgrade	*Signalling*	• ERTMS level 2 • No upgrade • No structures • No works in stations
Line category	Conventional	*Permanent way*	• No works in station
Design speed	161 – 200 km/h	*Earthworks*	
Environment	Rural	*Telecommunication*	• No structures
Terrain	Complex	*Electrification*	• No structures
Category of service	Mixed		
Axial load	6.4 – 7.2 ton/m		

The logics applied for the creation of the samples considered in the analysis of Tier 1 and 2 and of Tier 3 are reported below. The cost-drivers are presented following the same prioritisation considered in the REGIO Rail Unit Cost Tool (i.e. the last factor presented is the first excluded in case the identified sample does not encompass at least 40 projects).

<u>Sample for the analysis of Tier 1 and Tier 2</u>

The benchmark and forecasting analysis of the total, construction and base infrastructure unit cost is performed considering a cluster of project sharing the same **overall characteristics** of the project under assessment (left side of the table above, for the benchmarking analysis) or of the type of project for which the unit cost is investigated (forecasting function).

The REGIO Rail Unit Cost Tool filters the projects available in the data background and selects only those presenting the overall characteristics outlined in the table above. In case the sample includes less than 40 projects, the filter on the axial load is removed, to expand the selection. If the new sample still comprises less than 40 projects, the filter on the line category is neglected, and so on. The process is repeated until a sample of at least 40 projects is identified. <u>In any case, the minimum cost drivers considered are the type of works and the line category.</u>

<u>Sample for the analysis of Tier 3</u>

As detailed in Annex 9, the cost of each work included in the project is affected by factors which may differ from those affecting investment costs at whole project level. Some may well be the same, but some others may be different. E.g. it is of little significance (in terms of cost differences) whether electrification works are performed on a line that runs at more than 160 km/h or whether it reaches 200 km/h. Whereas the differentiation between the infrastructure categories remains significant as it impacts the characteristics (and costs) of the single project components. The detailed analysis of the factors impacting on the cost of each interventions/ element is reported in Annex 9.

The unit cost of signalling works considered in both the benchmarking and forecasting functions are then related to the unit cost of a sample of signalling interventions which:

- are performed on conventional lines;
- refer to the deployment of ERTMS level 2;
- do not relate exclusively to the upgrade of the ERTMS software;
- do not include the installation of specific signalling structure (e.g. radio-block centres) or the deployment of signalling station equipment.

The permanent way unit cost considered in both the benchmarking and forecasting functions are then related to the unit cost of a sample of interventions concerning the permanent way which:

- are performed on conventional lines;
- concern lines with a maximum design speed of 200 km/h;
- refer to mixed line (used by both passengers and freight trains);
- concern a railway line which axial load is comprised between 6.4 and 7.2 ton/m.

The earthwork unit cost considered in both the benchmarking and forecasting functions are then related to the unit cost of a sample of earthworks which:

- concern conventional lines;
- are carried out in rural environment;
- are carried out on a complex terrain.

The telecommunication unit cost considered in both the benchmarking and forecasting functions are then related to the unit cost of a sample of telecommunication interventions which:

- concern conventional lines;
- do not include the installation of specific telecommunication structure (e.g. radio towers).

The electrification unit cost considered in both the benchmarking and forecasting functions are then related to the unit cost of a sample of electrification interventions which:

- concern conventional lines;
- do not include the installation of specific electrification structures (e.g. traction substations).

5.1.2.3 PREDICTING FINAL COSTS

A detailed analysis has been performed to identify potential correlations between the expected costs and the final costs in rail infrastructure projects. The results, as presented in Figure 24, show an almost symmetrical bell shaped curve. Such analysis also clarifies that the final costs are not always higher, or lower compared to their estimates, as one might assume. On the other hand, a number of different factors which affect the estimated and final costs have been identified. Legislative and procurement procedures are among those. At a country level, these aspects can significantly vary in a short time period (e.g. due to reforms or introduction of new laws). Therefore, an historical analysis of the variation of estimated cost of investments at country level implicitly has some limitations.

Nonetheless, the analysis of the rail infrastructure market in the different MS reported in the country reports highlights some common trends that can be identified among Cohesion and non-Cohesion countries in terms of factors impacting on the final investment cost, e.g. level of detail of the planning process, procurement procedures and awarding criteria. Focusing the analysis on a dedicated cluster of countries - instead that on a single one - enables to reduce the impact of the changes of factors in a specific country.

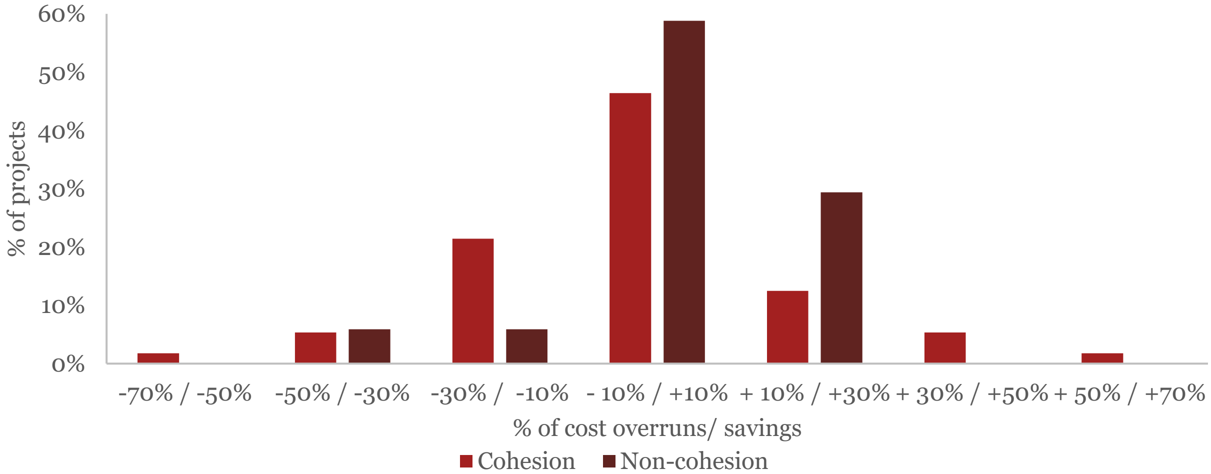
Therefore, to derive information on the expected variations of estimated cost, the distribution of the final cost included in the database as a percentage of estimated costs has been assessed considering the geographical factor.

Out of the ten projects with the highest cost savings percentages, **four projects are from the Czech Republic alone and eight are from Cohesion countries**. This finding for the Czech Republic corresponds to findings from the qualitative questionnaire where a respondent specified that the awarding criteria and competition market dynamics contribute strongly towards cost savings.

Particularly the impact of the procurement process is important because the cost estimate is often made before the tendering process is completed, wherefore bids lower than the estimated amount often cut the final project costs. Cohesion countries (in particular from Eastern Europe) have procurement awarding criteria centred on price. Therefore, the bids are often much lower than the cost estimates developed from the relevant national authority and can thus lead to savings. In the non-Cohesion countries on the other hand, procurement processes are mainly based on quality, wherefore cost savings will be rather rare provided the fact that it is likely that unforeseen cost factors drive up costs. Nonetheless, non-Cohesion countries tend to be more accurate when estimating costs and their estimations are based on more advanced stages of the planning process. As a result, despite more complex and frequent greenfield projects being executed in non-Cohesion MSs, the difference between final costs and estimated ones is overall lower, compared to Cohesion MS.

These findings are supported by the statistical results displayed in Figure 17, which shows that for non-Cohesion Countries there is a lower variance and less extreme values with both the median and the mean at 0% cost overruns, whereas for the Cohesion counterparts the Mean is at 3% cost overrun while the median is at 1% cost savings and the variance and outliers are larger.

Figure 24 - Distribution of cost overruns/ savings in Cohesion and non-Cohesion MSs



The analysis of final cost outlined that the factor impacting on the final cost of a rail infrastructure project is the geographical cluster (i.e. Cohesion or non-Cohesion country) where it is located.

This factor is taken into consideration in the analysis performed by the REGIO Rail Unit Cost Tool. During the analysis, the tool elaborates the distribution of the final cost of projects of the cluster identified based on the geographical cluster the project and calculates the probability that the cost variation falls within a specific range.

5.1.3 Construct & Implement

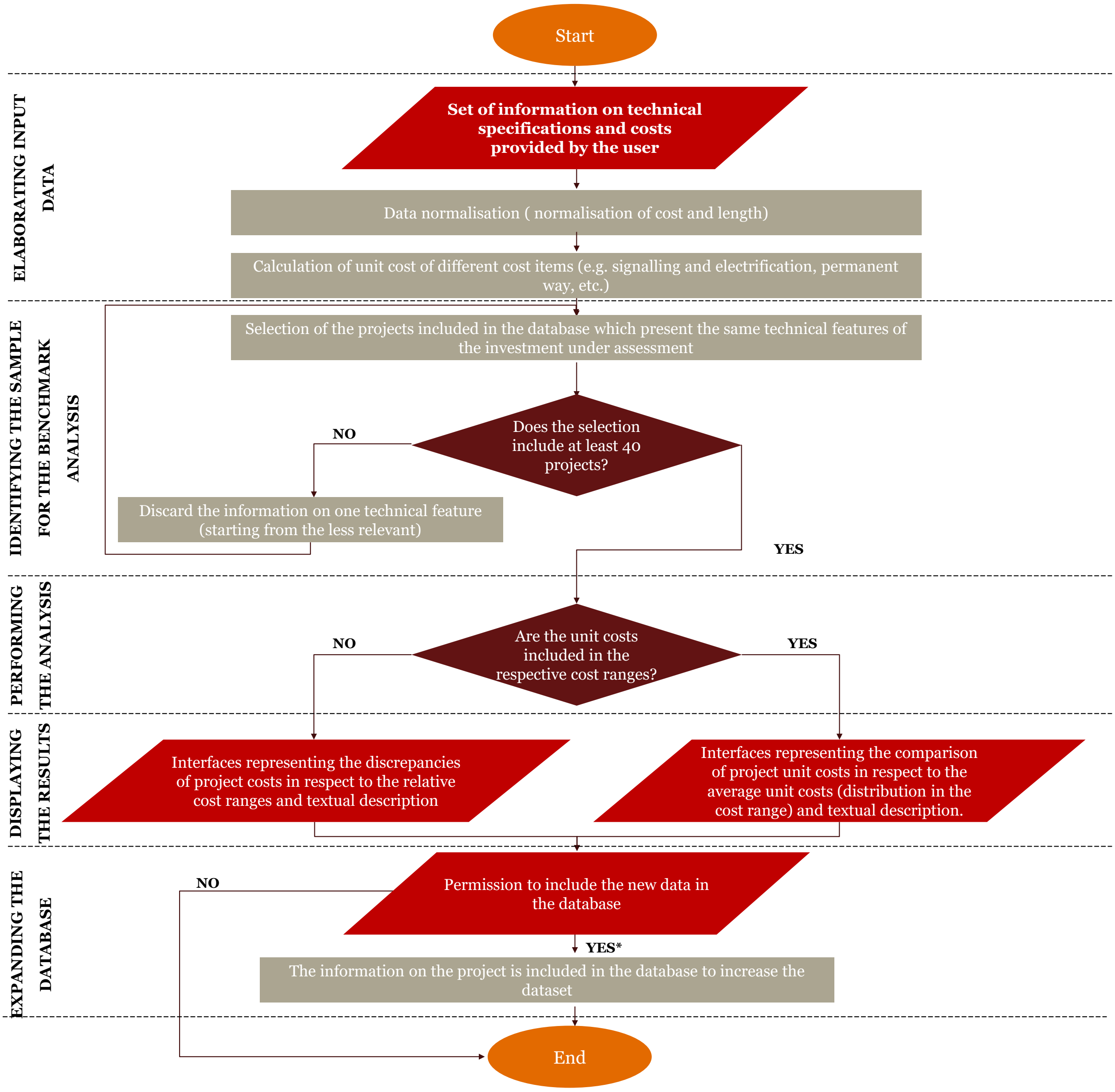
The REGIO Rail Unit Cost Tool enables the user to perform two different activities, for which similar logics are designed:

1. Benchmarking a project (i.e. insert the technical details and costs of a project not yet included in the database connected to the Tool);
2. Forecasting rail infrastructures cost (i.e. based on the comparable projects included in the database, derive median values for overall unit costs).

5.1.3.1 Benchmark

The definition of the logic of REGIO Rail Unit Cost Tool is functional to the creation of its algorithm, which is then used to construct the tool.

Figure 25 - The algorithm of the REGIO Rail Unit Cost Tool (benchmark)



The benchmakrking performed by the REGIO Rail Unit Cost Tool can be classified into five main phases:

- Elaborating input data;

- Identifying the sample for the benchmark analysis;
- Performing the benchmark analysis;
- Displaying the results;
- Expanding the database.

Elaborating input data. The REGIO Rail Unit Cost Tool is composed of several forms and interfaces to enable the user to input the information concerning the investment to evaluate. In order to ensure the validity of the outcomes, a minimum set of information must be input depending on the type of infrastructure work being assessed, defined on the basis of the validation of the analysis (see 5.1.4). The data on cost and length entered by the user are then normalised following the same approach adopted in the creation of the data background (see section 4.3). Unit cost are subsequently calculated for each cost item for which sufficient details have been provided.

Identifying the sample for the benchmark analysis. The REGIO Rail Unit Cost Tool applies the clustering logic and identifies the sample of projects to include in the benchmarking analysis (see section *Analysing uniform clusters of projects*).

Performing the benchmark analysis. The REGIO Rail Unit Cost Tool considers the unit cost of the projects included in the selected sample to define the unit cost distribution ranges used for the benchmark. Specifically,

- the minimum and the maximum values encountered in the sample correspond to the extreme values of the range;
- the other values are considered to calculate the quartiles, the median and the average of the distribution.

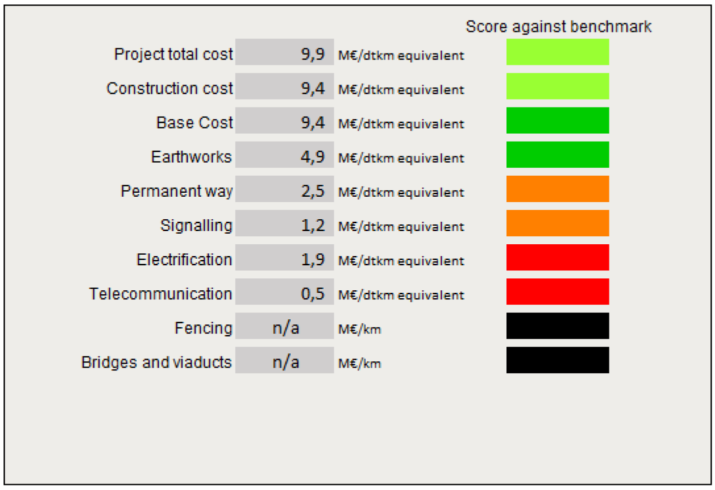
Once the cost range is defined, the REGIO Rail Unit Cost Tool assesses the position of the unit cost of the investment under assessment within the distribution.

The analysis is performed on different cost items, depending on the infrastructure category and the input provided by the user.

Displaying the results. The REGIO Rail Unit Cost Tool displays the results of the benchmarking analysis, showing the positioning of the unit cost of the investment assessed within the unit cost ranges derived from the projects included in the database.

1. **Summary overview.** A summary of the assessment is presented, enabling to immediately view the results of the benchmarking analysis (see Figure 19).

Figure 26 – Summary of the evaluation



In this case the user can find the numerical values of the unit cost per double-track kilometre of the line along with a five-colour scale that functions as a quick indicator as to whether the unit cost is in line with the average range of the benchmark for comparable projects.

Specifically:

- Green indicator: the unit cost of the investment is included within the average range of the benchmark;
- Light green indicator: the unit cost of the investment is in the range but lower than the average range of the benchmark;
- Orange indicator: the unit cost of the investment is in the range but is higher than the average range of the benchmark;
- Yellow indicator: the unit cost of the investment assessed falls outside the cost range derived from the sample of similar projects analysed, the value is lower than the minimum value
- Red indicator: the unit cost of the investment assessed falls outside the cost range derived from the sample of similar projects analysed, the value is higher than the maximum value
- Black indicator: in case no detail is provided to assess the project component against the benchmark, the coloured box would be black.

Additionally, an indication of the expected variation of estimated cost, based on the analysis of historical data is provided. It is possible to select the range to include a variation of ±10% as well as ±30% and have the tool providing the probability function of the observation to fall within the selected range.

5.1.3.2 INTERPRETING THE OUTCOMES OF THE EVALUATION

The results of the evaluation performed by the REGIO Rail Unit Cost Tool represent **the positioning of the unit cost of the project assessed within the distribution of the unit cost of similar projects.**

These results do not provide a measure of the quality of the project or of the cost estimates (e.g. a red-coloured indicator for the category signalling does not imply that the cost of the signalling system of the project has been poorly estimated).

An overview on the interpretation of the results of the analysis is provided below:

- a green-coloured indicator means that the unit cost of the project assessed is close to the median of the unit cost of comparable projects carried out in the past. Thus, the unit cost is close to its expected value (calculated considering the cost of previous projects);
- a light green-coloured indicator means that the unit cost of the project assessed is in the range of the unit cost of comparable projects carried out in the past. The unit cost is smaller than its expected value (calculated considering the cost of previous projects);
- an orange-coloured indicator means that the unit cost of the project assessed is in the range of the unit cost of comparable projects carried out in the past. The unit cost is higher than its expected value (calculated considering the cost of previous projects);
- a yellow-coloured indicator means that the unit cost of the investment evaluated is not comprised between the extreme values of the range and that the unit cost is minor to the minimum value of the range of the unit cost of comparable projects carried out in the past;
- a red-coloured indicator means that the unit cost of the investment evaluated is not comprised between the extreme values of the range and that the unit cost is major to the maximum value of the range of the unit cost of comparable projects carried out in the past;

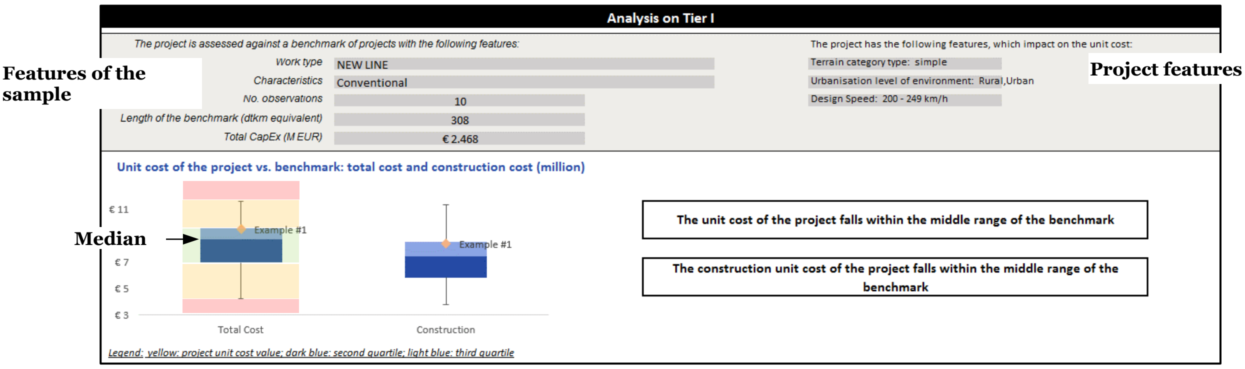
  The divergence may have different explanations. Firstly, the user should verify the values of the distribution and how far is the unit cost of the investment from the central share of the distribution. This way he/she can assess the magnitude of the divergence and consider if it can be considered negligible.

  Secondly, the user should check the characteristics of the sample used for the benchmark. To perform a reliable analysis, the REGIO Rail Unit Cost Tool may have disregarded some of the cost drivers that impact on the cost of the investment assessed (as outlined in section *Analysing uniform clusters of projects*). These may be the cause of the divergence.

Clicking on the buttons next to the summary overview, the user has the possibility to view the outcomes of the benchmark analysis for each tier.

2. **Analysis on Tier I.** The first level of analysis presents the results of the assessment of the unit total cost and unit construction cost. Two different charts display the positioning of the total and construction unit cost of the project under assessment within the related unit cost range derived from the data on the similar projects included in the database.

Figure 27 – Results of the assessment of total and construction unit cost

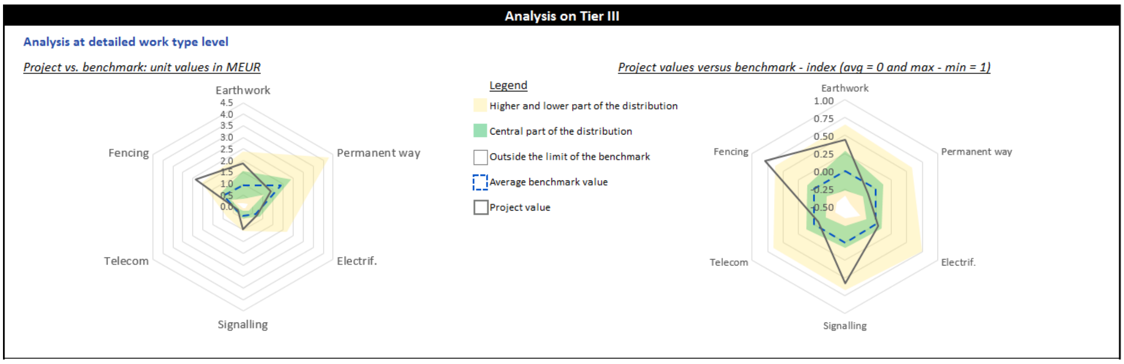


Would we wish to compare the summary results displayed in the *summary overview* ad the position of the project against the distribution of the benchmark shown in the figure for Tier I, we would consider that:

- The median of the distribution corresponds to the line between the dark and the light blue area.
- A green indicator in the summary overview corresponds to the unit cost of the project assessed falling within the dark or the light blue areas, which represent the two quartiles closer to the median value. Would the project fall in this range, it means that it is perfectly in line with the benchmark distribution.
- A yellow indicator is displayed if the unit cost of the project assessed falls within the first or the fourth quartiles of the distribution. In this case, the user should compare the features of the sample and the project features. Possibly, not all the factors that impact on the cost of the project assessed are considered in the benchmark analysis.
  To be noted, the graphical representation of the figure enables not only to identify whether the project under assessment lies. It also enables to understand how far the unit cost is from the average values of the benchmark. To this end, one project positioned to the far side of the range (but still within the range) would be differentiated from another similar one, which is still in the "yellow" area, but closer to the median.
- A red indicator in the summary overview corresponds to the unit cost of the project falling outside the cost range defined based on the unit cost of comparable projects included in the data background. It may be necessary to further investigate the reasons behind the out-of-range value, requiring information to the project promoter.

3. **Analysis on Tier II.** The second level of the analysis concerns the base infrastructure cost. The same chart used for the Tier 1 shows the positioning of the base infrastructure unit cost of the project assessed within the range derived from the base infrastructure unit cost of the projects of the sample.
4. **Analysis on Tier III.** The third level of the analysis presents the results of the assessment of the components of the base infrastructure cost, in case the relative data on to the investment under assessment is provided by the user. The outcome of the analysis is summarised in two "spider charts", presenting the positioning of the unit cost of the different elements within the respective cost ranges (see figure below).

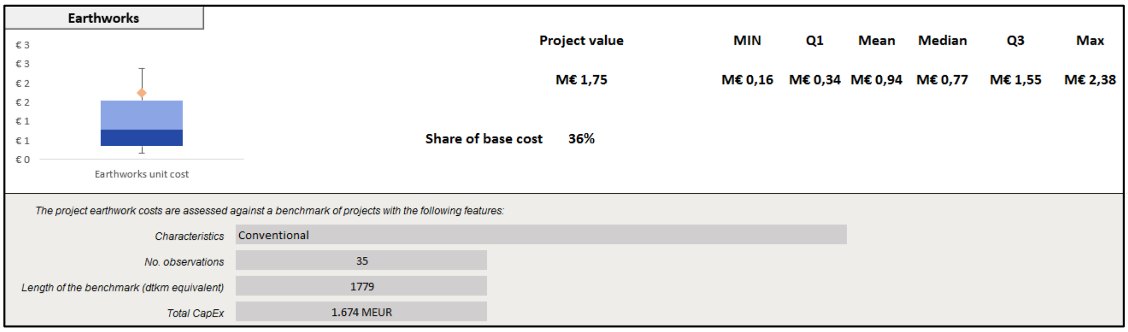
Figure 28 – Summary of the analysis on Tier III



The first chart is based on the unit cost values expressed in M EUR. The cost ranges for the different elements usually vary (e.g. on average the unit of the permanent way is considerably higher than the unit cost of fencing). Therefore, the scale of the chart reaches higher values for some categories. In the second chart, the same scale is used for all the cost items. This enables to view the results of the assessment of each categories together. In this chart, the average values correspond to zero. Negative variations indicate that the unit cost of the project assessed is lower than the average unit cost of the projects included in the sample. Positive values correspond to unit costs higher than the average. The colours of the charts facilitate the identification of the quartiles of the distribution and correspond to the colour of the indicator included in the summary overview (e.g. if the unit cost of the project falls within the third quartile – in the green area – the corresponding indicator in the summary overview will be green-coloured).

The outcomes of the analysis of each cost element is also represented in details using the same chart presented in the analysis of Tier I and Tier II. The figure below presents the detailed results of the assessment of earthworks. As mentioned above, the sample used for the analysis on Tier III is different from that considered for the other Tiers. In the section presenting the detailed assessment of each element of the Tier III, the information on the sample considered in the benchmark is provided.

Figure 29 – Example of detailed analysis on Tier III



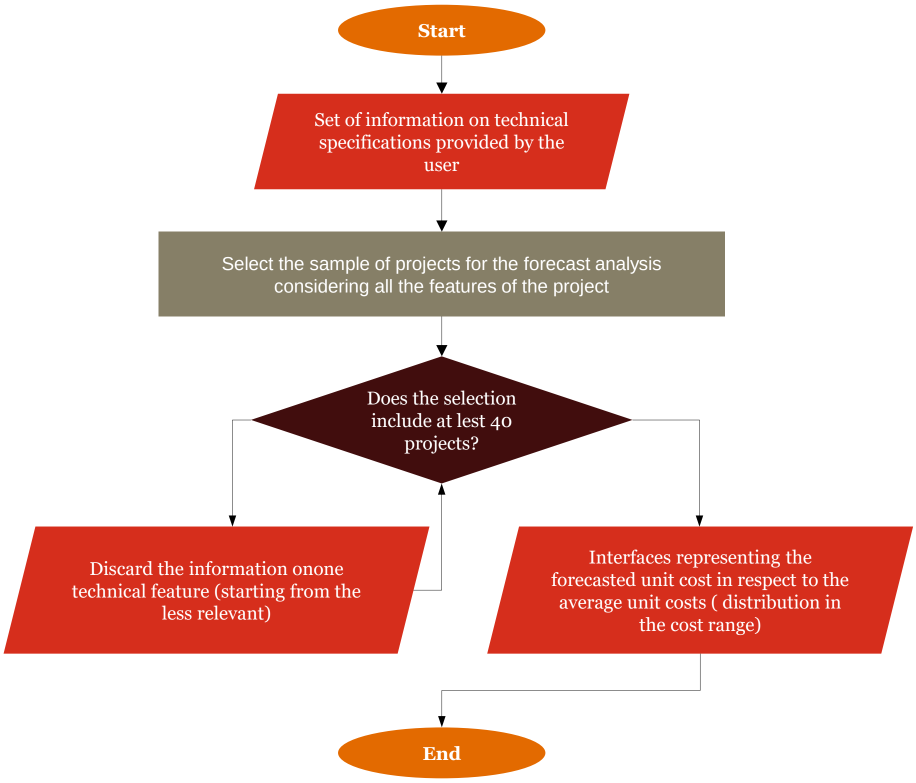
Expanding the database. Once the benchmarking analysis is completed, the user has the opportunity to include data related to the investment assessed in the database. **Therefore, the database can be expanded and the cost ranges resulting from the benchmark analysis can change each time the REGIO Rail Unit Cost Tool is used.** Furthermore, the availability of additional data will enable to increase the accuracy of the clustering and therefore the reliability of the results.

The possibility to add new projects to the database applies also in case the investment assessed results to fall outside the cost ranges identified. Indeed, this does not necessarily mean that the cost of the investment is not appropriate to its characteristics and/or is wrongly estimated. It may be the case that the project presents specific characteristics which justifies the higher/ lower cost.

5.1.3.3 FORECAST

The REGIO Rail Unit Cost Tool algorithm follows a logics similar to that of the Benchmarking.

Figure 30 - The algorithm of the REGIO Rail Unit Cost Tool (forecast)

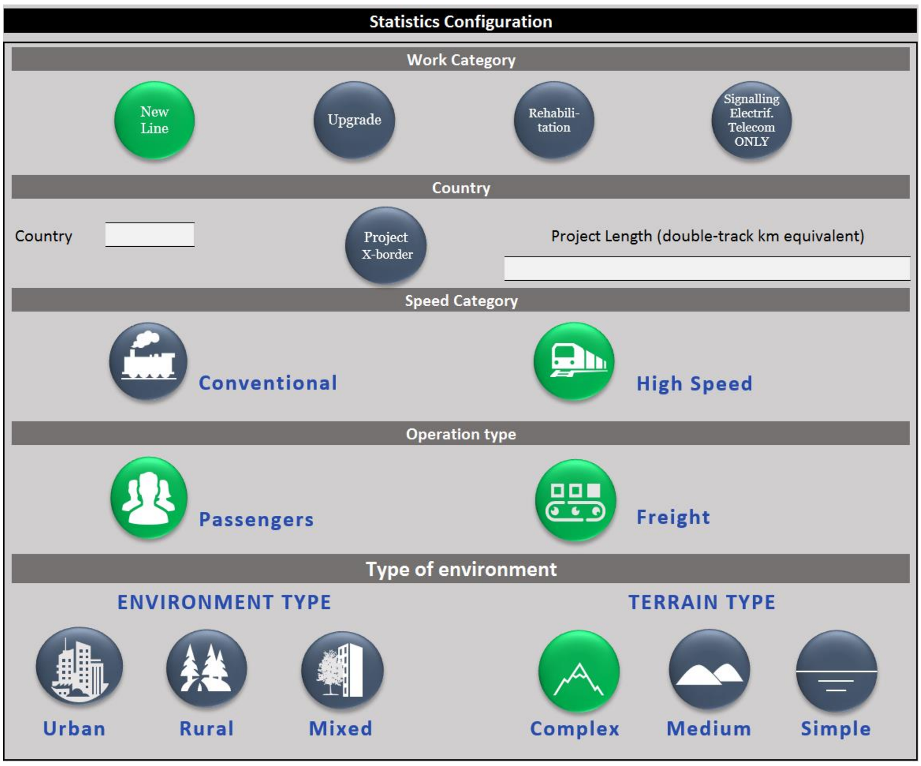


Four actions are undertaken:

1. Elaborating input data;
2. Identifying the sample for the forecasting;
3. Performing the forecasting;
4. Displaying the results.

Elaborating input data. Users, to develop the forecast, need to input the technical characteristic of the project. An input mask facilitates the user to input the information related to the technical characteristic of the project. Users can select the characteristic of the project by clicking on the button correspondent. In order to ensure the validity of the outcomes, the Tool does not start the analysis if the user has not inserted a minimum set of information and if there are incoherencies within the information provided.

Figure 31 - Input configuration forecast



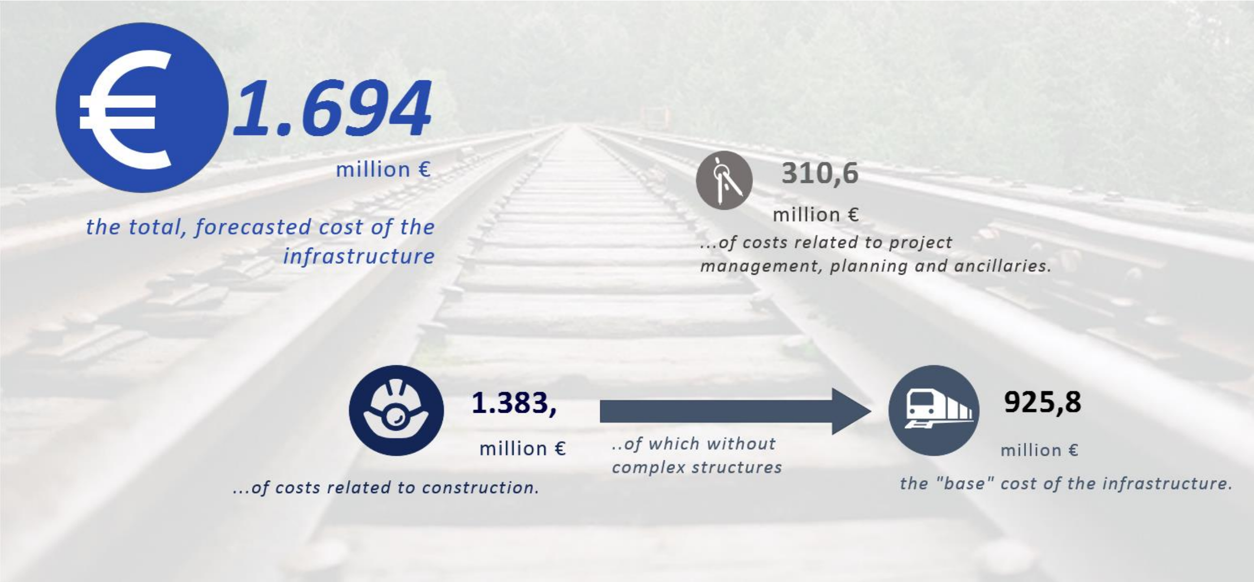
Identifying the sample for the forecasting. The REGIO Rail Unit Cost Tool applies the clustering logic and identifies the sample of projects to perform the forecasting. The smallest sample accepted for the analysis is 40 projects. In case the database does not include more than 40 projects with the characteristics chosen by the user, the tool discards the information on one technical feature until it reaches a significant sample. This is done for the type of environment and operation-type information, while the minimum categorisation (work category, speed category) is maintained regardless the sample size. The sample creation is described in more details in the sub-section "Design" (5.1.2).

Performing the forecasting. The REGIO Rail Unit Cost Tool considers the unit cost of the projects included in the sample selected to define the expected unit cost ranges. The forecast is made for all cost items, from Tier 1 to Tier 3, providing an increasing level of detail.

Displaying the results. The REGIO Rail Unit Cost Tool select the cost information of projects with the characteristic provided by the user and displays the results of the forecast.

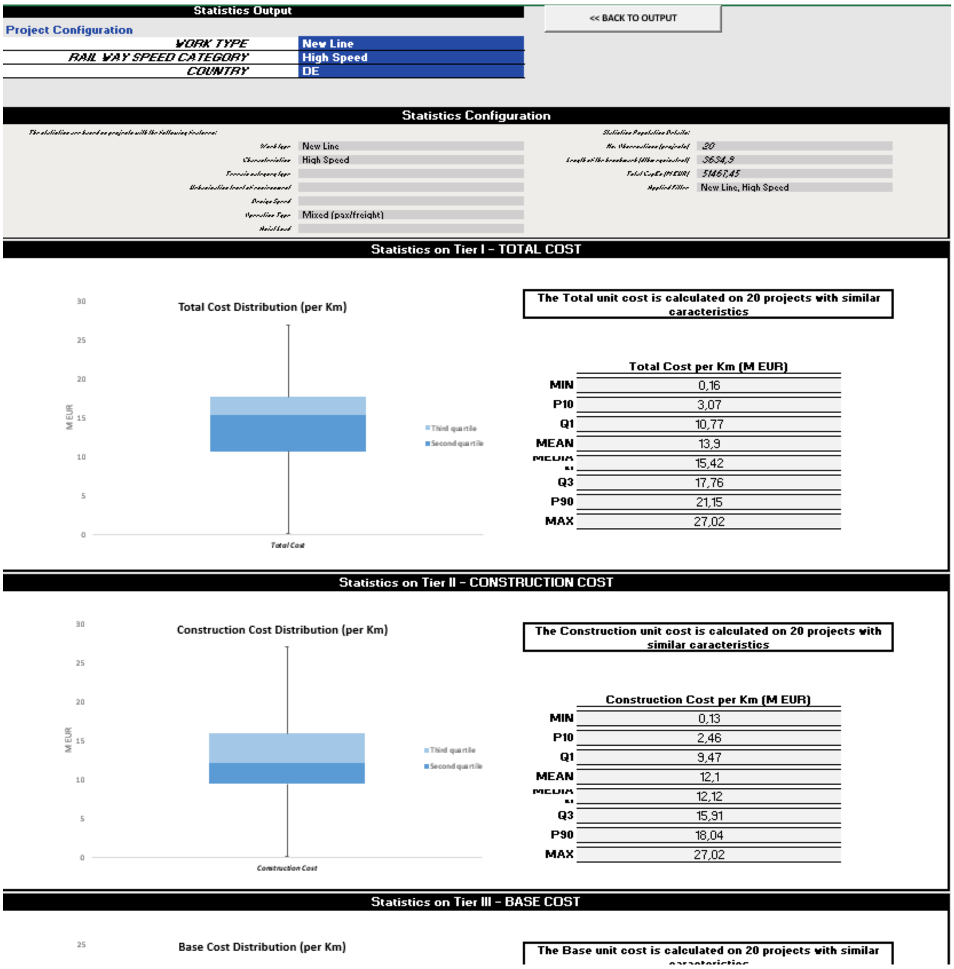
1. **Forecast summary.** The tool provides an overview of the forecasted cost of the infrastructure. The information provided are: total forecasted cost, project management, planning and ancillaries cost, construction forecasted cost and base forecasted cost.

Figure 32 Forecasting general results



1. **Statistical details.** The information displayed present the same detail of the benchmarking results. Forecasted cost are provided for each cost element, from Tier 1 to Tier 3. The Tool displays for each cost element the distribution of the total cost of the projects in the database with the same technical characteristic and a boxplot graph, which graphically represent the information provided.

Figure 33 Statistical details



5.1.4 Validate

The validation phase aims to verify the validity of the approach undertaken and the possibility to identify cost ranges that enable the REGIO Rail Unit Cost Tool to perform the benchmark and forecast analysis.

As mentioned above, it should be considered that REGIO Rail Unit Cost Tool is a "**self-learning tool**", therefore while the approach can be validated, the cost ranges are expected to improve as the tool is used.

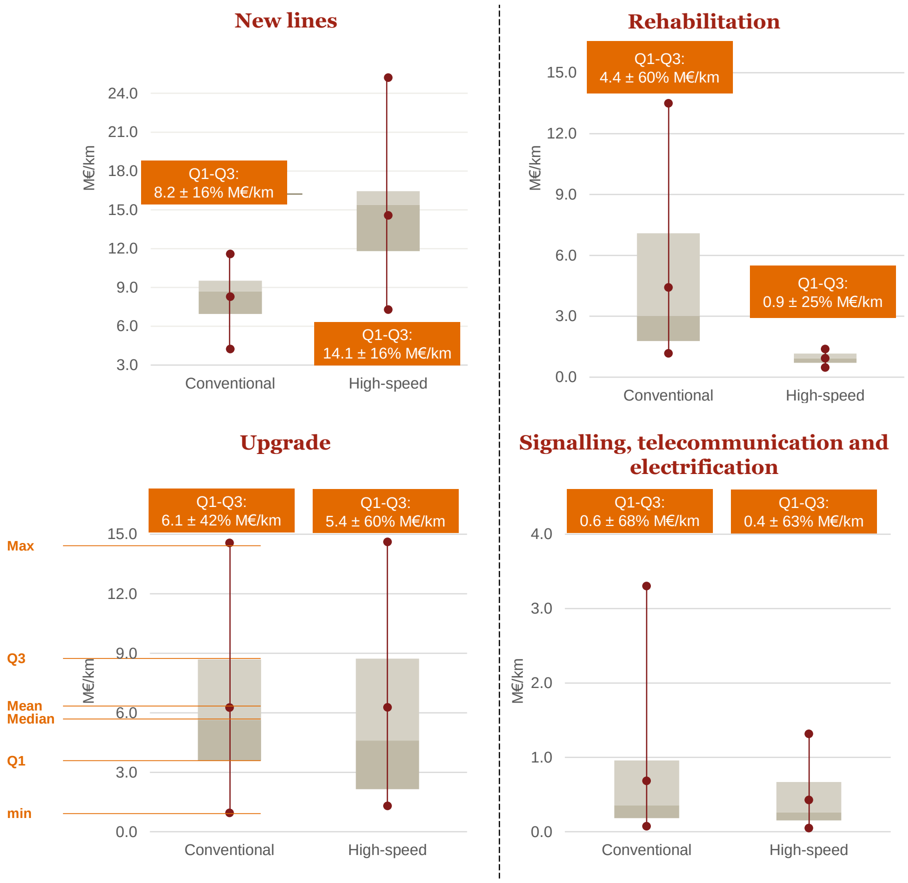
THE COST RANGES OF INFRASTRUCTURE CATEGORIES

The analysis of total investment cost

The first level on analysis performed to validate the approach concerns the general costs of the infrastructure.

The validity of the approach adopted for the analysis is confirmed by the good results achieved even considering general cost. Indeed, the distance between maximum and minimum values of the unit total cost is lower than that that encountered in the majority of the previous studies (see 2), most likely as a result of the accurate classification activity carried out in building up the database.

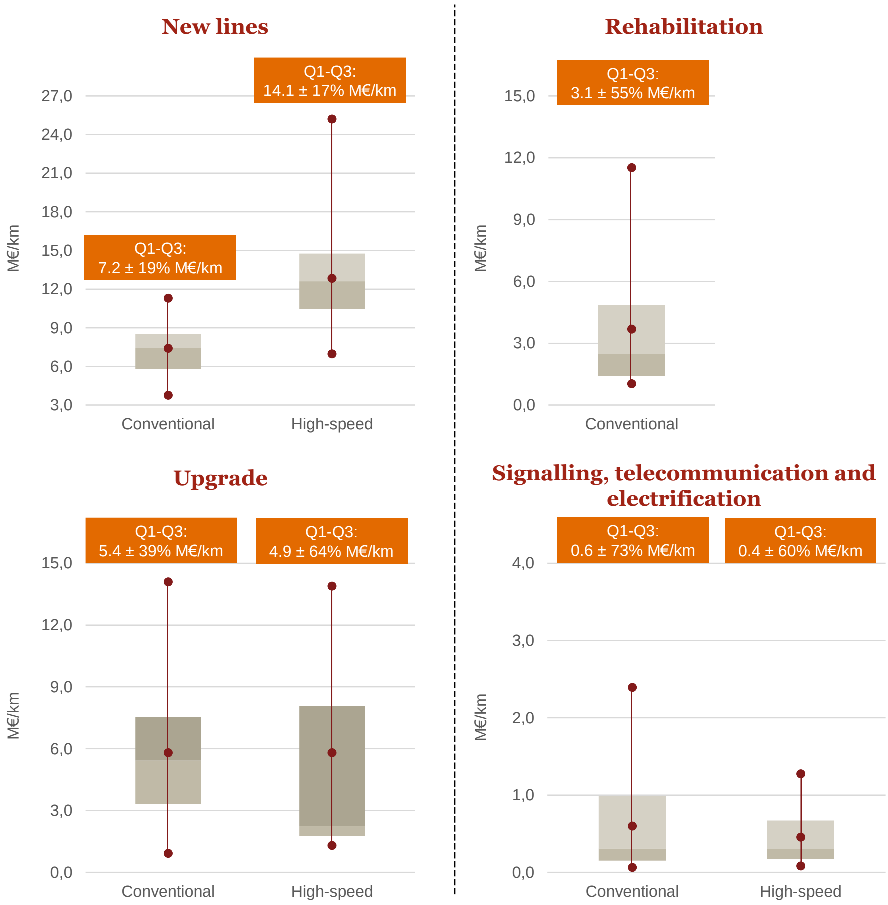
Figure 34 - Total investment unit cost ranges per infrastructure category



In particular, the unit cost range for new lines can be already considered a valid reference to assess a new investment. The categories "Rehabilitation", "Upgrade", "Signalling, telecommunication and electrification", which require several infrastructure interventions, present a larger cost range. Yet these can be considered useful to perform a preliminary benchmark or forecasting. The only exception is represented by the rehabilitation of high-speed, which is the category with the narrowest cost range due to the limited number of observations. This, in turn, is attributed to the low number of interventions carried out to restore the capacity or the conditions of high-speed lines, which are generally recently constructed infrastructure.

The analysis of the construction unit cost shows how, refining the scope of the analysis, more accurate cost ranges can be achieved, thus improving the reliability of the benchmark and forecasting exercise.

Figure 35 - Construction unit cost ranges per infrastructure category

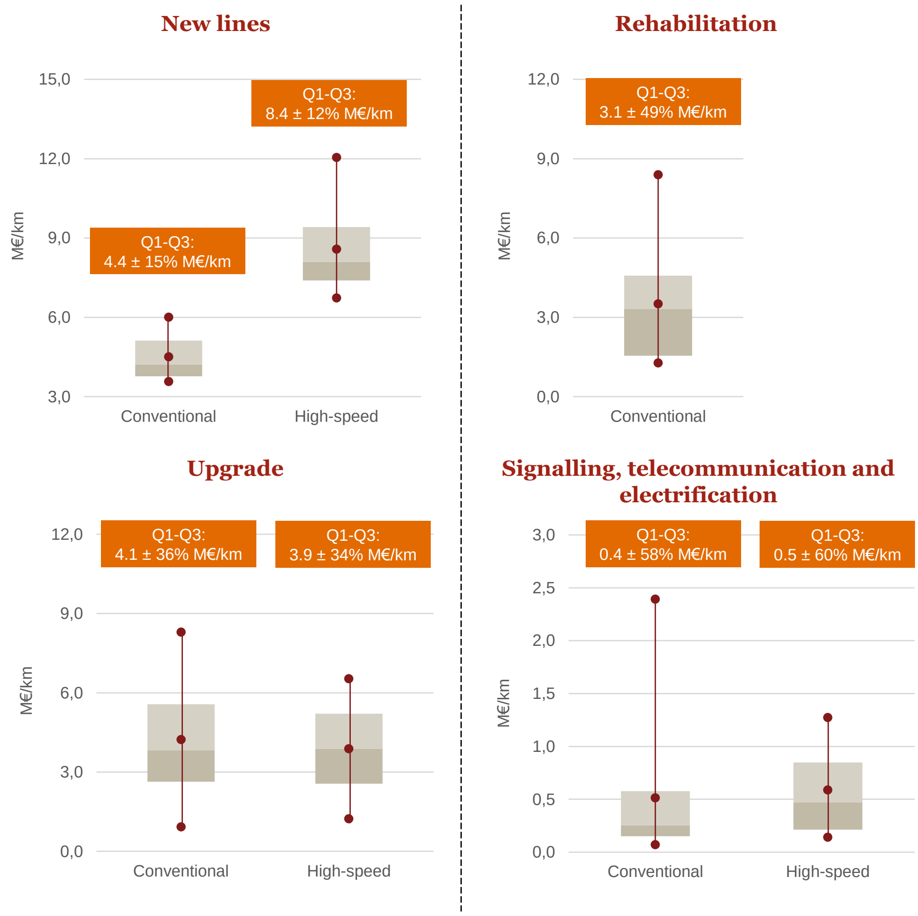


A lower variability is in fact encountered in the case of upgrade interventions on conventional lines. In case of high speed lines, the analysis shows that the variability of the cost range is related to the presence of quite different projects within the sample, some of them significantly on the minimum value of the range (as demonstrated by the proximity of the median and minimum values). The need of shifting the analysis to cost items in order to perform a significant analysis is therefore confirmed for the high speed lines, as well as for the other category such as signalling, electrification and telecommunication.

The analysis of the base infrastructure cost

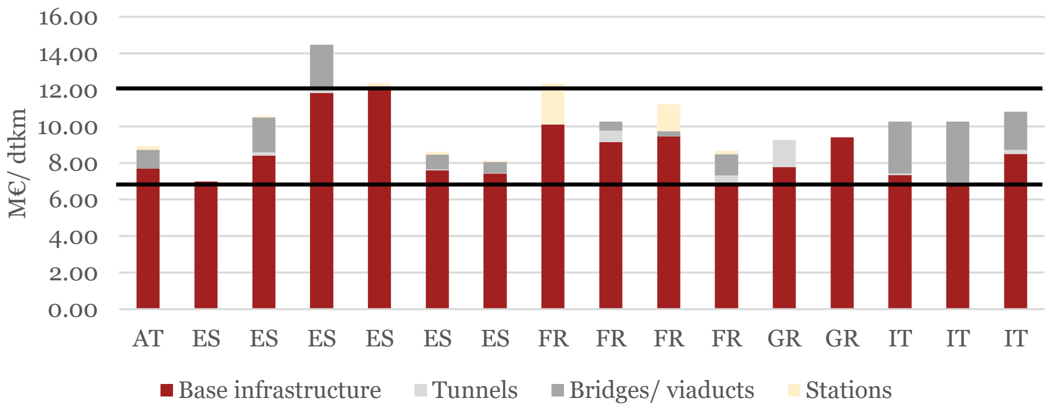
The analysis of the base infrastructure unit cost enables to reduce overall the variability of the cost ranges, confirming that applying a bottom-up approach on an accurately clustered dataset leads to the identification of accurate unit cost ranges to be considered in a benchmark and forecasting analysis.

Figure 36 – Base infrastructure unit cost ranges per infrastructure category



Particularly, the analysis of the base infrastructure cost leads to the identification of especially narrow cost ranges for investments aiming to the deployment of new lines. Indeed, shifting the analysis to the base infrastructure cost enables to exclude the cost of civil structures and stations which have a relevant impact on the variability of the construction unit costs. Once they are subtracted from the construction costs, it is possible to observe how unit cost ranges become closer.

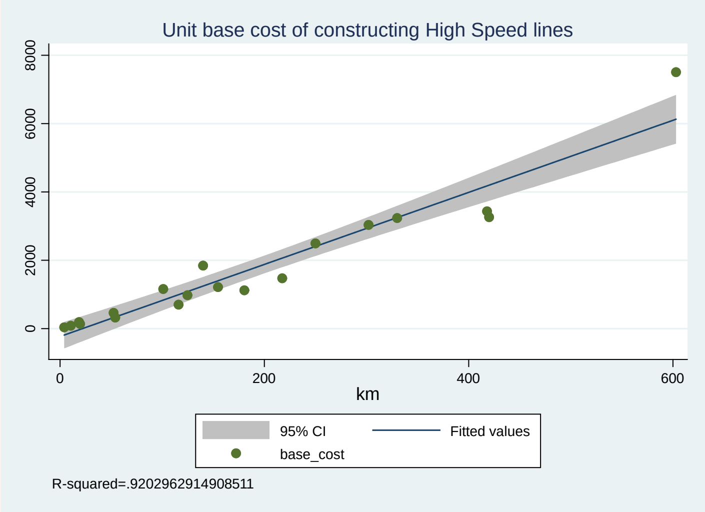
Figure 37 – Base infrastructure costs of new high-speed lines



It is worth mentioning that, concerning high-speed lines, the **unit costs of the base infrastructure have been calculated on a sample representing approximately 45% of the high-speed railway lines in use in the EU.**

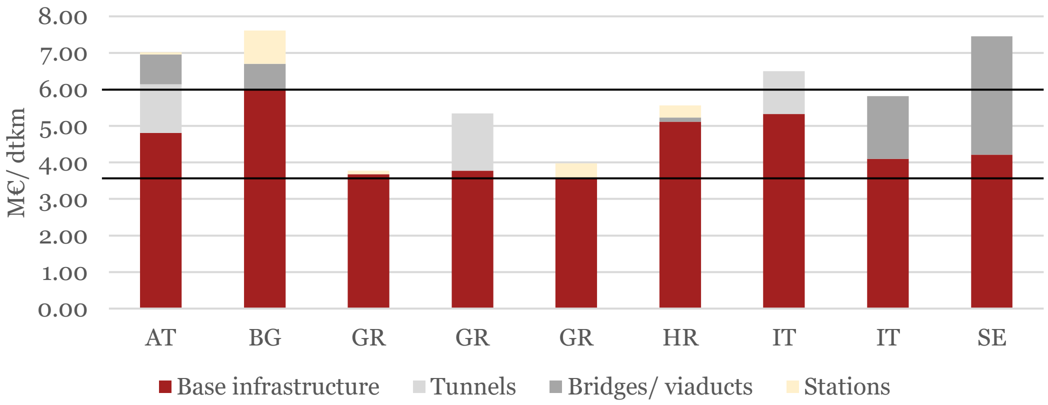
As outlined in figure below, a clear correlation between the infrastructure cost and railway length can be established.

Figure 38 - correlation between the length and the base cost of new HS lines, 95% Confidence Interval



The impact of structures and station on the construction cost of new line investments is confirmed by the analysis of the projects concerning conventional line. Also in this case, a unit cost range is identified once share related to big structures and stations is subtracted from construction cost.

Figure 39 – Base infrastructure costs of new conventional lines



Assessment of the results

The analysis of the base infrastructure unit cost enables to identify a narrow cost range which enables to perform an accurate benchmark or forecast analysis of investments referring to a new line.

Therefore, the base infrastructure cost is the minimum level of detail required for the evaluation of this work type category.

As for the categories “rehabilitation”, “upgrade”, “signalling, electrification and telecommunication” the analysis of the total investment and the base infrastructure cost enables to identify general cost ranges. Nonetheless, given the diversity of the works encompassed by this kind of projects, an accurate analysis, in particular for the benchmarking, of the investment cost can be performed only if the analysis of unit cost is carried out on each of the single interventions included in the investment (e.g.

signalling, electrification, earthworks, etc.).

As a result, the information about the cost of each intervention should be available in order to enable the carrying out of a detailed assessment of the investment cost at tier three level.

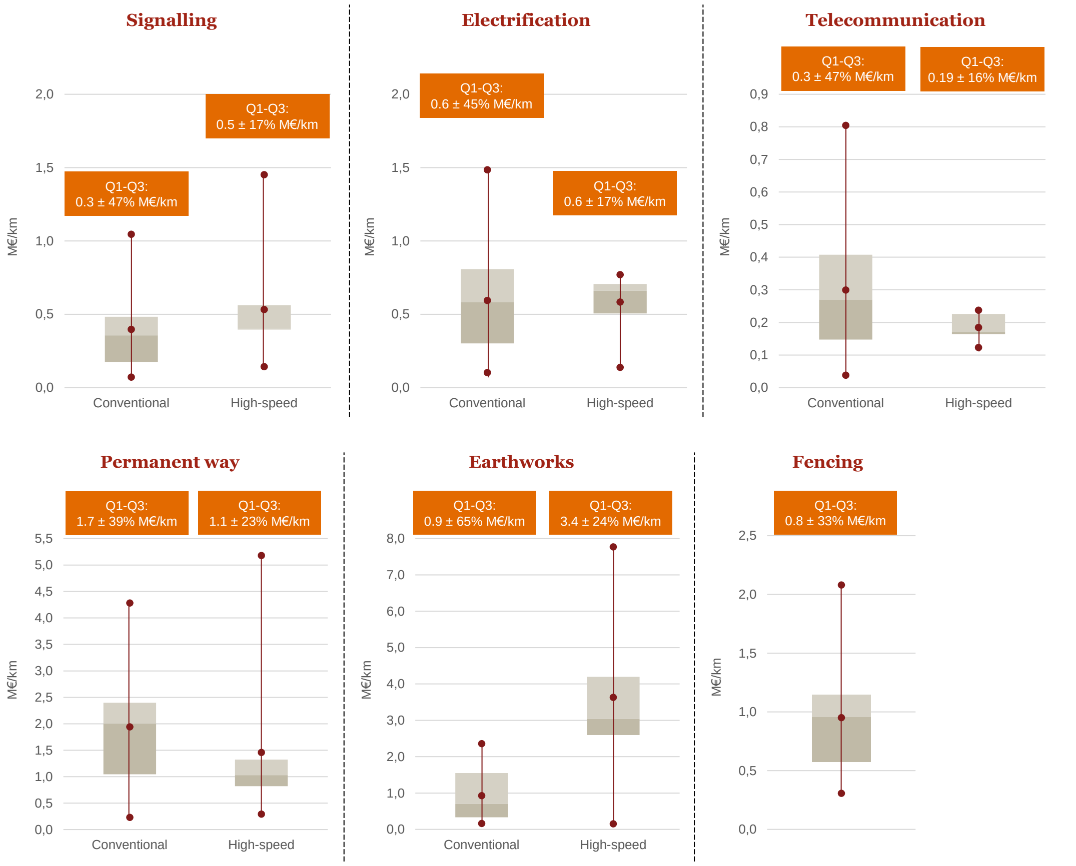
The analysis of the single cost items

The last level of analysis the REGIO Rail Unit Cost Tool is designed to assess single cost-items.

In this case, while differentiating between conventional and high-speed lines is still necessary (the infrastructure category impacts the technical specifications thus the characteristics of the components analysed), discriminating between the work type categories becomes less relevant. It is not necessary in fact to cluster projects based on the similarity of the interventions encompassed, being each interventions/ cost item analysed separately.

It is worth noting that fencing is the only component that remains the same for both conventional and high-speed railways as its technical characteristics do not depend on the infrastructure category.

Figure 40 – Unit cost of single components



The analysis of the single components enables to identify for each of them cost ranges that can be used to perform a reliable analysis. A higher variability is observed for the high speed category with regards to the cost of permanent way and earthworks. However, this is likely due to difficulties in isolating the cost related to these two categories for the projects analysed. As soon as new data is included, the cost range will be refined.

Assessment of the results

The analysis of the single cost components enables to identify cost ranges that can be used for a benchmarking (and potentially forecasting) analysis for each of them. These will further be refined as new data is included in the REGIO Rail Unit Cost Tool.

The unit cost of the rehabilitation, upgrade, signalling, electrification and telecommunication interventions can be properly assessed if the analysis is carried out on each intervention include in the investment.

As a result, the information about the cost of each intervention should be available in order to enable the carrying out of a detailed assessment of the investment cost at tier three level.

From a merely statistical perspective, it would be possible that the unit cost of all the works included in a project lie within the respective cost-ranges yet close to the minimum (or maximum) values.

These extreme cases would present acceptable scores for each single component, albeit the total investment and the base infrastructure unit cost would be poorly evaluated. The difference among the levels of acceptance at the different tiers is relative to the fact that the frequencies of projects that have combined levels of very low (or very high) unit costs for all tier three works are almost irrelevant, statistically. Coherently, they are not close to the average base infrastructure unit cost.

The limited statistical relevance of these extreme values was confirmed by the analysis of combined tier three unit cost frequencies, which also support the statistical relevance of the analysis of tier two.

The frequency simulation analysis

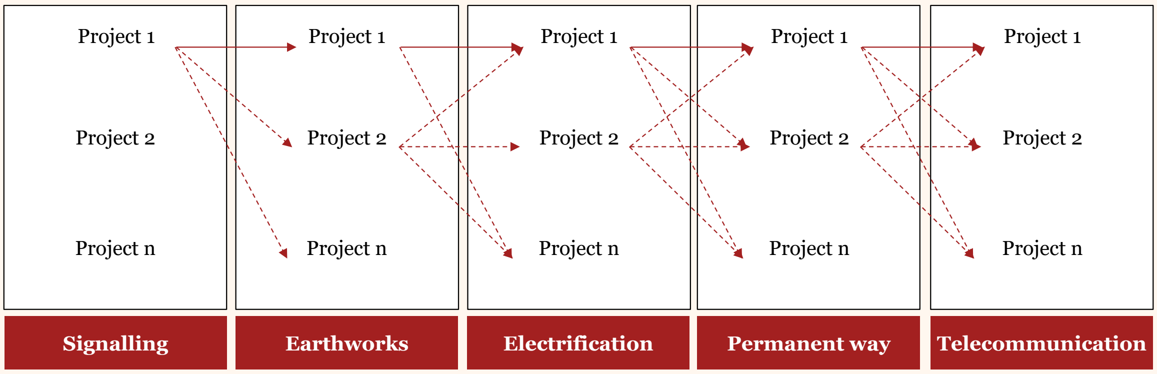
The frequency simulation analysis simulates all possible cost combinations of each single cost elements considered in the tier three from all projects available. The reasoning behind this analysis is that the individual cost items (excluding fencing, as this is not always necessary) make up the base infrastructure cost, thus the two cost-ranges should be comparable[35].

The analysis was performed on a database including all values for the individual cost elements included in the dataset– each cost element separated from the project it belonged to. Therefore, if for a project all of those five elements were available, the five elements were treated individually.

All possible different combinations of the individual cost elements was created, similar to a probability tree using the statistical tool MATLAB (e.g. the cheapest elements of each category added up make up the lowest potential unit cost, while the most expensive elements of each category make up the highest potential unit cost and all possibilities in between – view figure below). This way it was possible to simulate a set of **290,049,750 unique combinations for conventional rail projects.**

[35] Due to a lack of detailed data, it is not possible to take into consideration the cost of minor hydraulic structures and interfaces, which are comprised in the base infrastructure cost. Therefore, a slight variation of the cost is already anticipated.

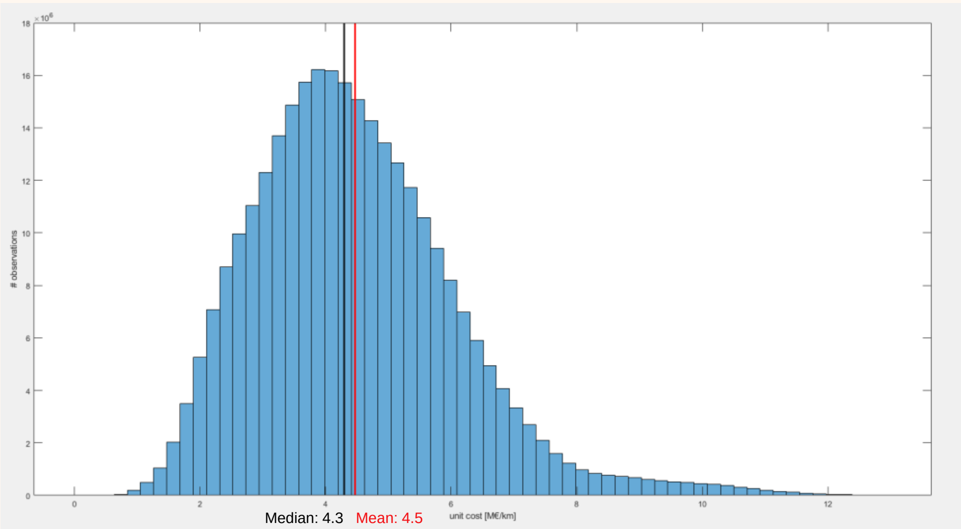
Figure 41 - Methodological approach of combining different elements from different projects



The distribution of the total unit costs is a simulation of different hypothetical project costs and therefore the real distribution of potential base unit costs, based on the assumption from the distribution and the frequency of occurrences of costs, an approximation of the probability distribution of the unit cost range for real projects can be derived.

Furthermore, it is also possible to simulate average costs because the hypothetical scenarios themselves represent potential projects that combined allow for the analysis of average as well as medium values of the simulation sample.

Figure 42 –Frequencies of the combined tier-three unit cost



The analysis confirms that the probability that a project with the extreme values of each cost element is encountered is significantly low. **The most likely values are completely in line with the base infrastructure unit cost range, thus such analysis further validates the results of the work carried out.**

5.1.5 Operate

This section defines the three main components of the REGIO Rail Unit Cost Tool: **people, process** and **technology.**

PEOPLE

The REGIO Rail Unit Cost Tool is designed as a tool to support decision making process concerning rail infrastructure investments projects.

Specifically, the current version of the REGIO Rail Unit Cost Tool is intended to be used by EC commissioners, officers and decision-makers to benchmark the unit cost of a rail infrastructure investment to be supported by EU funds with the cost of similar projects carried out over the EU. Based on the results of the assessment, the decision-makers may require further information which explain unexpectedly high or low investment costs. In addition, the tool enables to estimate the cost of the infrastructure based on main characteristic of the project. The information that the tool provides depend on the number of projects available in the database.

In this regard, an instruction guide has been developed with the objective to support the end users of the tool. The guide presents several sections and provide instructions on how to insert information into the REGIO Rail Unit Cost Tool and presents a set of definitions of the terms used to undertake the benchmark and forecasting analyses. For further details please consult the REGIO Rail Unit Cost Tool instruction guide annexed to this report.

REGIO Rail Unit Cost Tool instructions

A specific document has been developed to support users of the REGIO Rail Unit Cost Tool, each section dealing with a phases of the use of the tool (see Annex 12).

Chapter of the user guide:

1. Introduction;
2. The structure of the REGIO Rail Unit Cost Tool user guide;
3. Home page;
4. Data input;
5. Guide to output understanding.

PROCESS

The functionalities of the REGIO Rail Unit Cost Tool are designed focusing on the process related to the assessment of rail infrastructure project co-funded by EU funds.

The application concerning these projects is submitted for approval by regional and national MS to DG REGIO. This is the stage where the REGIO Rail Unit Cost Tool is deemed useful. The officer/ commissioner who receives the application has in fact the possibility to input the project data into the REGIO Rail Unit Cost Tool and compare the investment cost with that of similar projects.

The process is organised in four macro stages:

- Data Input;
- Benchmarking/ Forecasting analysis;
- Dataset expansion;
- Maintenance and update.

Data Input

Data Input is the first stage of the process. The user will input the information of the investment project that is being analysed. The guide will play a key role during this phase, supporting the users in the correct input entering.

As previously mentioned, the REGIO Rail Unit Cost Tool is a living tool and therefore is constantly updated as a result of the inclusion of new projects in the database. It is therefore vital for its correct functioning, to ensure that the information input is coherent and is correct. If projects with wrong cost or technical specifications is included in the database, then the benchmark analysis might be invalidated and lead to unrealistic results.

For the forecasting analysis it is only required to input the features that characterise the infrastructure type for which it is researched to estimate the cost.

Benchmarking/ Forecasting analysis

Once all the inputs have been inserted – according to the type of analysis to be undertaken – the REGIO Rail Unit Cost Tool opens a new sheet with the results of the benchmark for the respective project or of the forecast, depending on the function activated. The tool applies the clustering logic and identifies the sample of projects to perform the analysis on. The results of the analysis are presented through specific charts (as outlined in the instruction guide).

Expansion of the database

One of the most useful functionality of the REGIO Rail Unit Cost Tool is that it offers the user the possibility to include the assessed project into the database for future analyses, in case the project is deemed representative. Therefore, the database will expand over time as more projects are inserted by the users. However it is necessary that the information added to the database are correct and coherent, otherwise the risk of creating unbalances in the set parameters might become very high.

Maintenance and update

The maintenance and update of the REGIO Rail Unit Cost Tool is necessary to ensure its correct functioning in the following years. With regards to the maintenance is vital to ensure that the projects included in the database are relevant and correctly input.

A yearly update of the tool is required to ensure that it is able to function properly. Prices must be normalised to the base year. The REGIO Rail Unit Cost Tool currently converts all the costs of the different European states to a common PPP and are adapted to 2016 to allow to create a European benchmark tool. The costs that are included in the future require to be translated accordingly, applying conversion parameters. To ensure that the tool can be easily updated, the sources of the factors are selected in order to be: *public, freely accessible* and *from well recognised institutions*. All data can be gathered from Eurostat and OECD, as specified in Annex 10.

As the database will continuously expand with new projects being included over time, a periodic revision of the data is suggested, to ensure that there has been no mistake in inputting projects. The analysis has to be performed directly in the tool database, which is accessible to the product owner after he/she unlocks the file using the correct password (detailed instructions on the maintenance of the database are provided in Annex 10).

TECHNOLOGY

The current version of the REGIO Rail Unit Cost Tool is ready to be introduced within the process to assess rail infrastructure investment. It is currently based on a Visual Basic macro which operates on excel.

System technical requirements

The REGIO Rail Unit Cost Tool has been developed and tested using systems with the following requirements:

- Operating system: Windows 8;
- Software: Microsoft Excel 2010 or later versions;
- VBA libraries:
    - Microsoft Office 15 Object Library 15 or later versions;
    - Microsoft Excel 15 Object Library 15 or later versions.

This does not mean that the REGIO Rail Unit Cost Tool is not compatible with systems with different features, but rather that the tool full functionality is not ensured on these systems.

The different interfaces are the different worksheets of the same file. The REGIO Rail Unit Cost Tool is therefore, light, easily accessible and ready to use. A series of error checks have been included to support the user with the correct insertion of the values as well as ensuring a coherent dataset. However the technology undertaken also presents some limits, in particular the database behind the REGIO Rail Unit Cost Tool is included in the excel file itself, therefore in case of duplication of file, the database will be duplicated too. Such duplication might lead to differences in terms of databases (and therefore results of the benchmark and forecast analysis) as different projects might be included in the database. To overcome this limitation, it would be necessary to transfer the REGIO Rail Unit Cost Tool from a Microsoft Office Excel based tool to a software or web-base. This option would also enable to expand the user pool, e.g. it would be possible to grant access to the tool to Managing Authorities to input data on rail infrastructure investments through a digital form.

6 Conclusions and recommendations

6.1 Conclusions of the study

Investments in railway are at the core of the transport investment strategy of the Union, enabling to shift traffic from less environmentally friendly and less sustainable transport modes to rail. While in other sectors the technological advancements enabled an easier collection and spread of data, in the rail sector this process has not occurred. Access to information on investment projects data and related studies is yet more limited in the sector compared to other transport modes. The ability to analyse and process the data at European level has long been researched by different academic authors as well as by European Institutions. The aim was to enable more accurate considerations when evaluating projects either at national programme level and project level when providing support to co-funding. This study was therefore prepared with the objective to contribute in:

- establishing a common understanding on factors influencing rail infrastructure investment cost,
- elaborating a methodological approach to identify and calculate unit cost in rail infrastructure investments,
- providing input to decision-makers for high-level evaluation of the costs of rail infrastructure programmes and projects,
- provide forecast on the cost of rail infrastructure
- assisting authorities in charge of rail infrastructure programmes and projects to communicate effectively to a wider set of stakeholders about project cost.

The authors of the study performed a thorough review of an extraordinary amount of documents related to rail infrastructure projects developed in Europe since late 1990s. All project documentation that both the stakeholders and the Commission provided was gathered, organised and analysed in detail. This resulted in hundreds of projects being processed and included in a database. Previous studies have attempted comparable analyses. Nonetheless, they were focused on a smaller scale or on a lower level of detail. This study is therefore something new in the Industry and enabled the authors to observe a number of insights, which are summarised in the following short paragraphs.

The Database

Following the literature review, and the stakeholder consultation activities, the authors of the study developed a structured project database, which stored information on technical features and costs of rail infrastructure projects. The database was initially built based on the data made available by DG REGIO, INEA, European IM and PwC's contacts on over 860 projects. The project screening and selection process, led to the finalisation of the database containing over 200 rail infrastructure projects from all over Europe. The database itself is a relevant outcome of the study, given its size and level of accuracy.

Such project mapping was functional to the development of the REGIO Rail Unit Cost Tool, the digital tool designed to support the assessment of investment cost of rail infrastructure project based on a benchmark with the investment cost of comparable projects.

The data collection phase has only been a starting point. It lasted several months and involved different teams in different countries working side by side with Infrastructure Managers and local authorities to collect data stored sometimes on paper, in the local language. The effort resulted in a vast amount of data collected and processed. Nevertheless, once the REGIO Rail Unit Cost Tool is properly used, the database will be fed with new data at the same rate at which new projects are evaluated. This way, the database could increase significantly in a rather limited amount of time.

The REGIO Rail Unit Cost Tool

The main outcome of the study is the REGIO Rail Unit Cost Tool: a tool that enables to store information on projects and to statistically analyse costs and features to provide benchmarks or cost statistics on the basis of projects presenting similar characteristics. The main features of the REGIO Rail Unit Cost Tool are its user friendliness and the ability to adapt throughout time to an increasing dataset. The tool indeed automatically adapts the size and features to those of the project being assessed. The more projects are assessed and added to the database that the tool uses as source, the more the tool is able to be accurate. This as a consequence of increasing the population of the distribution which is used as reference as well as the more accurate the selection of the said population.

When used to benchmark projects, the REGIO Rail Unit Cost Tool is designed to simplify the comparison of one project with a wide sample of other rail projects, regardless the time and location. The tool indeed normalises the costs of the investments to enable the comparison across the dimension of time and geographical locations. The perfect normalisation cannot be achieved, anyway. Differences across countries in terms of rail sector practices, procurement laws, competition, safety requirements, etc. can only partially be explained by the normalisation factors. Country-level analyses are produced to support the analysis of projects, together with the outcomes of the REGIO Rail Unit Cost Tool. At the same time, comparing single projects with a panel of geographically and, to a certain extent, technically different projects contributes to avoiding assessing observations biased due to local peculiarities against similarly biased observations.

All considered, the REGIO Rail Unit Cost Tool is designed to aid in the identification of potential discrepancies at project level as well as at project component level, which are to be further investigated.

The REGIO Rail Unit Cost Tool is a benchmarking tool; in addition to the benchmarking function, **the tool can use the database of information to support the forecasting of project costs** based on the distribution values of the costs of projects with similar characteristics.

While most observations are expected over time to be more and more distributed close to the median values, the following must be always considered:

- the REGIO Rail Unit Cost Tool does not produce technical assessments of projects; it compares projects technically and financially. Would a project under assessment lie outside the range identified by the REGIO Rail Unit Cost Tool for the specific type of work and features, it would not necessarily mean that the project is wrongly designed, its costs are wrongly estimated, nor that the REGIO Rail Unit Cost Tool itself is mistaken. Peculiarities can occur in all sectors, rail investment one included; such cases would be likely to be accounted for as outliers, in a probability distribution, without affecting the soundness of the project itself nor of the tool evaluating it. The REGIO Rail Unit Cost Tool is designed to produce likely suggestions on what could potentially be a project that requires closer attention.
- The REGIO Rail Unit Cost Tool is dependent on the information that is stored in its database. Would the database be populated with odd observations, the results would be equally odd. It is therefore paramount to use the REGIO Rail Unit Cost Tool following the indications provided in the attached guide and using the tool the way it is designed and developed to.

Differences across Member States investing in rail

Investments across the EU tend to vary depending on the geographical and development level of the country. While in Cohesion countries rail investments tend to focus on renovating and upgrading under-maintained and obsolete railway networks; in more modern Member States, the focus is often set on developing new, high-speed networks.

The different investment focus widens the gap that in any case stands between Cohesion and non-Cohesion member states in terms of procurement procedures, composition of the market and project planning practice (see below).

Regardless the country, all railway networks are investing in renewing the signalling system, to comply with the EU Regulations that require ERTMS to be installed on the TEN-T Core and Comprehensive network.

Final costs and planned costs

While it was the first intention to design specular analyses for estimated and actual costs, data availability has not been sufficient to do it. Information on final costs has indeed been made available only for a fraction of projects analysed. Nevertheless the analysis of the data collected as well as the information gathered through the stakeholder consultation enabled to identify common patterns across Cohesion countries and across non-Cohesion countries in terms of elements creating gaps between estimated costs and final costs. In particular:

- The level of detail in planning projects tends to be higher in non-Cohesion countries. The more detailed planning contributes to keep the final costs closer to the planned costs, compared to countries where projects are detailed less.
- In the case of new technologies being implemented, the lack of experience in estimating costs often leads to significant discrepancies between planned and final costs. For example, this has been the case for some of the first ERTMS projects.
- Different procurement practices, as well as different market configurations create significant changes in terms of variations between projects planned costs and final costs. The European Commission developed a constant monitor of the public procurement practices in Member States, rating their efficiency and effectiveness. The same considerations apply for rail investment procurement practices.

These elements combined matched with an overall distribution of frequencies of cost overruns and cost savings, which has been included in the REGIO Rail Unit Cost Tool, to provide an estimate of the cost variation that can be expected, based on historical data. While this result presents a certain validity, also in consideration of the sample used, it does not match project-level characteristics with the risk of final costs being largely different from the planned ones. Nor it is able to take into account the different political, geographical and regulatory elements that may have impacted the projects with financial risks. This risk assessment indeed requires a level of detail, which goes beyond the reach of this study. It is suggested that this element be investigated further in the future (see recommendations below).

6.2 Recommendations

During the elaboration of the present study, the authors have identified a number of elements, which the Commission may consider to pursue in the future, to enhance the achievement of their stated objectives in the rail sector. These have been summarised in the recommendations listed below. This list proposes specific topics and actions to be considered by the Commission in terms of both amendments of current practices and further topics to assess through new studies.

Reporting Requirements

Although good and promising results were achieved, the current study demonstrated that there are large data gaps when analysing the unit cost that prevent a more detailed assessment, particularly when analysing costs on a Member State level and per different types of infrastructure. Those data gaps arise as a consequence of the lack of consistent technical specifications and costs for the majority of projects. Such information is not available to the Commission either, and the data collection of such specific data elements is often a challenging activity, even for the Infrastructure Managers themselves, because there are no requirements to provide and store this data in a harmonised format. Furthermore, provided that the REGIO Rail Unit Cost Tool is designed to evolve and improve the accuracy of the unit cost analysis over time with new user inputs, it would be useful to align the information requested in the application forms with the information required to perform the cost analysis as per the REGIO Rail Unit Cost Tool. Therefore, ***it might be beneficial to include those elements in the reporting formalities required for project applications of the new programming period.*** A proposal for the forms which may be used to gather information on new projects is provided in Annex 16.

Required level of planning

The qualitative analysis of the study has shown that there are large differences between the planning processes of the infrastructure projects among different Member States. There is a particularly severe difference in the level of detail that is required by different countries. In a lot of countries, the cost estimates are done on either a general project level or using a lump sum instead of detailed cost breakdowns per project element. This leads to less accuracy of the cost estimates and a larger discrepancy between the cost estimates and final costs. As a consequence ***it could be beneficial to implement a planning requirement for projects applying to EU funds from all Member States that consists of a standard, common level.***

Expanding the current scope of the study

Although this study's analysis on final costs compared to estimates yields significant results based on the data presently available to the study, it is to note that there are many limitations provided the small amount of data available given the different scope of the study. Therefore, conducting a separate study focused on such issues would be beneficial and it would allow to further investigate the correlation between cost overruns and external factors as well as to examine project based risk factors and cost contingencies. The REGIO Rail Unit Cost Tool has been developed to support the assessment of the investment cost in rail infrastructure projects, provide indications on the expected final cost of the project and estimate the cost of rail infrastructure. However it is not designed to assess project risks. Notably, although increased construction time, unforeseen geographical barriers, mistakes in or poor design, etc. are all elements that increase investment costs, it is not possible to develop a statistically relevant sample. Therefore, it would be ideal ***to conduct an additional study focussing on the identification, understanding and analysis of project risks*** which will be able to provide recommendations on how to tackle incidents encountered during the development of a railway infrastructure.

Upgrading the REGIO Rail Unit Cost Tool

As previously mentioned the REGIO Rail Unit Cost Tool is currently based on a Visual Basic macro and it operates on Microsoft Office Excel. The programme is therefore light, easily accessible and ready to use. However such architecture presents some limitations as the database behind the REGIO Rail Unit Cost Tool is included in the Excel file itself. In case of duplication of file, the database will be duplicated too. Currently the tool has been developed for a specific end user, DG REGIO, and therefore such limitation is expected to have a limited impact. In the future however, the tool might be shared with other stakeholders involved in railway infrastructure development, namely MS Authorities, Infrastructure Managers, Railway Undertakings, other DGs, etc. Therefore, ***it would be necessary to develop an upgrade of the current tool, linking a unique database (web or software based) with the different versions of the tool itself.*** Such upgrade would allow to increase the usability and transferability of the tool, which could be used in different location, updating a unique and common database.

HOW TO OBTAIN EU PUBLICATIONS

Free publications:

- one copy:
  via EU Bookshop (http://bookshop.europa.eu);

- more than one copy or posters/maps:
  from the European Union's representations (http://ec.europa.eu/represent_en.htm);
  from the delegations in non-EU countries (http://eeas.europa.eu/delegations/index_en.htm);
  by contacting the Europe Direct service (http://europa.eu/europedirect/index_en.htm) or calling 00 800 6 7 8 9 10 11 (freephone number from anywhere in the EU) (*).

  (*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

Priced publications:

- via EU Bookshop (http://bookshop.europa.eu).

Priced subscriptions:

- via one of the sales agents of the Publications Office of the European Union (http://publications.europa.eu/others/agents/index_en.htm).

KN-01-18-846-EN-N

Publications Office

doi: 10.2776/296711