Addressees	**Kyriacos Ktenas (DG REGIO – F.1)**
Author	**PwC**
Project Name	**Assessment of unit costs (standard prices) of rail projects (Capital Expenses)**
Subject	**PROGRESS NOTE 20 June 2017**

1.	**State of play and key progress from the inception meeting (May 10th) to June 20th**

	Data collection
	• Creation of the database has been completed, according to the classification of information suggested in the Inception Report. The Database will be used to store quantitative data on projects. Minor changes to the structure of the database might happen during the processing and disaggregation of data.
	
	• Preparation of the stakeholders consultation package and submission of the survey:
	
	 o We have finalised the Stakeholders contact list that is composed of 3 main groups: Managing Authorities (45 MA, 17 out of which at National level and 25 at Regional level), rail Infrastructure Managers (35, including IM of cross-border projects), Ministries of Transport/ Agencies (38 in total).
	
	 o Differently to what planned in the Inception Report, due to the limited time at disposal, we decided to deliver the qualitative and quantitative consultation at the same time.
	
	 o The Stakeholders consultation package was confirmed by the EC on 7th of June 2017, what signifies the start of the stakeholders consultation process.
	
	 o **The qualitative consultation** was sent to national and regional Managing Authorities (MA), to the relevant Ministries and Agencies(MIA) in all 26 MS. The stated deadline for data collection is 30 June 2017. The consultation package to MA and MIA is composed of:
	
	  ▪ a cover letter in the text of the email
	  ▪ the online questionnaire (tailor-made for each category of stakeholder):
	  ▪ the comfort letter from the EC.
	
	 o **The qualitative/ quantitative consultation** was sent to the IM in all 26 MS. The stated deadline for data collection is 14 July 2017. The consultation package to IM is composed of:
	
	  ▪ a cover letter in the text of the email
	  ▪ the online qualitative questionnaire
	  ▪ the comfort letter from the EC,
	  ▪ the instructions for the quantitative survey,
	  ▪ the list of projects for which we request info,
	  ▪ the excel file for the collection of project info.
	
	• Status of Stakeholders consultation
	
	 o On 9 June 2017 we sent the stakeholders consultation package to all the stakeholders (281 people consulted).
	
	 o Stakeholders' commitment to the process is of utmost importance and PwC is putting maximum commitment in order to receive as high as possible response rate and data coverage. Thus, starting from June 12th the team has being performing

follow up calls to:

- verify the stakeholders contact list correctness

- establish personal contacts

- explain the purpose of the study and sensibiliser on the importance of the stakeholders engagement

- offer support for quantitative data collection. Considering the particular importance of certain IM in terms of number of projects we have offered our full availability to support them in performing the task.

o To monitor the status of the stakeholders consultation a monitoring sheet was prepared. The monitoring sheet will be shared with the EC on a weekly basis. The first monitoring sheet is attached to this report.

Data processing, disaggregation and database compiling

o In the reporting period the team analysed the documents received from DG REGIO and INEA.

o From May 10th to May 20th, we set-up a uniform approach to ensure consistency to the data extracted by different people and we trained the data collection team. The team is composed of people from the PwC transportation team. We mobilised the network where needed (e.g. Poland).

o From the week starting on May 22nd, the team has been analysing information available to extract required data for each project. The data cleaning exercise and disaggregation exercise is done on project-by-project basis, while reviewing the paper documents collected.

o Data are processed, disaggregated to the extent it is possible and inserted in the database. No assumption is made. On average, a person is able to finalise 2 to 4 projects per day.

o Until now, we have analysed 90 projects (out of 541). As a preliminary overview, considering the projects analysed until now, it can be said that the level of data availability is relatively high for the Tier 1 level, but for the Tier 2 and Tier 3 very low availability of data can be stated.

2.	**Issues and challenges**

We decided not to use data from the DG REGIO database (WP 10). We have performed a detail checks on the sample of 5 projects comparing the information inserted in the DG REGIO database with the one reported in the project fiches, contracts, decision, and in 3 out of 5 cases that we checked, we found inconsistencies/lack of accuracy as follows:

- **2002PL16PPT016: Modernisation of E30 railway line**
  o Total Cost [M€]: 90,499 includes "Expenditures incurred before application" (Planning/design fees = 2,458 M€) as reported in the Commission Decision (draft). In the WP10 DB, the total estimated costs (**1_Total_e**) are 90,499 M€ (correctly), but the estimated Planning_and_design cost (**1_Planing_and_desing_e**) cell is empty. In synthesis, the disaggregated information, although available, has not been reported in the DB;
  o The Contingencies reported in the Commission Decision (draft) are 7,257 M€. The contingencies estimated (**1_Contingency_e**) in the WP10 DB are 57,48 M€. The figure reported in the DB is not correct.

- **2002CZ16PPT013: Optimisation of Zabreh na Morave - Krasikov railway section**
  - o The Total eligible cost reported in the Commission Decision of 17/10/2002 (121,299 M€) has been filled in the DB as total actual cost (**1_Total_e**=121,30 M€). <u>The information is correct, but the classification is wrong.</u>
  - o The contingencies estimated in the application form are 10,97 M€, while the estimated contingency on the DB (**1_Contingency_e**) is 0,01 M€. <u>The data reported in the DB is wrong.</u>
- **2000HU16PPT001: Rehabilitation of the Budapest-Cegléd-Szolnok-Lőkösháza railway line**
  - o The total estimated cost reported in the Commission Decision of 21/12/2000 report a total eligible cost of 126,00 M€ and 0,85 M€ of expenditures incurred before the application (of which 0,701 M€ are for "Planning/design fee" and 0,149 M€ are for "Technical Assistance"). The total estimated cost (**1_Total_e**) reported in the DB is 126,85 (correctly as the sum of both eligible costs and expenditures incurred before the application), but the both the estimated planning and design cost (**1_Planing_and_desing_e**) and estimated technical assistance cost (**1_Technical_assistance_e**) have not been reported in the DB. <u>In synthesis, the disaggregated information, although available, has not been reported in the DB.</u>
  - o The estimated project length in the database (**U_Project_length_e**=256 kms) refers to the length of "rails new and rehabilitated" (in the progress report of the 31/12/2008) and not to the Length of the Railway Line rehabilitation 2*78 kms (in the progress report of the 31/12/2008). <u>The data reported in the DB is wrong.</u>

Since we cannot fully rely on the data presented in the DG REGIO WP 10 database, we need to make our own data elaboration from all available documents from the DG REGIO Database what makes the exercise more time and energy consuming than expected at the beginning.

Due to the high number of projects and the need to screen all papers to extract data and to involve a large number of stakeholders, the data collection phase as a whole might take more time than planned at the beginning. As far as:

- the data collection from DG REGIO and INEA documents is concerned, we plan to report data for half of the projects in the list by the delivery of the first progress report.

- The collection of qualitative data from stakeholders, we plan to get most of the questionnaires filled in by the delivery of the first progress report

- The collection of quantitative data from stakeholders, even if deadline is July 14th, stakeholders started asking for extension of the deadline. For the time being, we are not able to plan when we will get information.

ADIF informed us today that they do not have any time to collect the requested information. Our contact will discuss with the management the possibility to accept our offer to provide them support.

We kindly ask your support to engage ADIF. Our contact is Mr Jaro Lorenzo (Deputy Director of Programming and Investment Evaluation) ljaro@adif.es, +34 915 067 186

3.	**Next steps**
	• We are continuing stakeholders consultation process. Follow up calls and e-mails will continuously be sent to ensure participation from all 3 categories of the stakeholders with

the highest emphasis on the IM.

- On June 26th we will inform the EC on the need for support to stimulate reaction from specific stakeholders (especially IM), if needed.

- In case of a need, a bilateral meeting with the stakeholders will be organised. In particular, on purpose visit/ conf call or video conf call to specific MS will be organised when:

  - IM ask for our support to work on requested quantitative information

  - based on the answers to qualitative questionnaires we found interesting comments or we require additional information

  - we need to collect information on case studies

  - we do not have enough information on a specific Country coming from projects.

- Data processing, disaggregation and database compiling will continue.

- As soon as we have compiled the database with information for half of the projects (indicatively by the mid of July), we will start exploring methodologies for processing information as well as suggesting possible case studies.

- The First Progress report will be delivered by second week of July 2017, as agreed with the Commission.

Addressees	**Kyriacos Ktenas (DG REGIO – F.1)**
Author	**PwC**
Project Name	**Assessment of unit costs (standard prices) of rail projects (Capital Expenses)**
Subject	**PROGRESS NOTE 07 July 2017**

1. | **State of play and key progress from the last Progress Note (June 20th) to July 7th**

Data collection - Stakeholders consultations

- The status of the stakeholders consultation process is synthetized in the following tables.

	Invitation Emails sent (No. of people) 09/06/2017	*Reminder emails sent (No. of people) 21/06/2017 - 23/06/2017*	*First Follow-up call established (No. of people) 15/06/2017 - 20/06/2017*	*Second Follow-up calls established (No. of people) 21/06/2017 - 04/07/2017*
MA	52	35	7	33
IM	95	72	44	38
MoT	94	69	2	33
Total	**241**	**176**	**53**	**104**

	Number of organisations invited	*Number of organisations responded to qualitative survey (as of 4 July 2017)*	*Current response ratio to qualitative survey (organisations) as of 4 July 2017*	*No. of organisations responded to quantitative survey (as of 4 July 2017)*	*Total No. of qualitative and quantitative responses as of 4 July 2017*	*Organisations unwilling/unable to respond to survey*
MA	42	7	17%	n/a	3	2
IM	38	9	23%	2	10	2
MoT	40	13	35%	n/a	6	0
Total	**120**	**29**	**24%**	**2 (0%)**	**19**	**4**

- Main outcomes of the follow-up calls:
  - Most of the countries replied that they were looking for the most suitable respondent to provide us with the requested data and that they will work on it (7 MA, 18 IM and 11 MoT, see below for details)
  - Austria - asked for a 1 week extension. Agreed to send the qualitative data by 3rd week of July.
  - Belgium - organised a VC with the Belgium IM Infrabel NV on the 6th of July . Asked for the extension of the deadline for the quantitative part. The qualitative questionnaire will most likely be delivered by the end of July.
  - Finland - IM will provide us with the quantitative data required by the 14 June. On 5 July was performed another call to update us on the progress.
  - France - several messages were dropped in the voice mail (in order to obtain news from them). We are constantly recalling and monitoring the contact with this IM.

Until now, we have not established a contact with them.

- o Greece - reported difficulties in filling the qualitative part of the survey. We provided them with additional guidance and reminded that the task is not just the qualitative survey, but also the quantitative data collection. They forwarded a request for the quantitative data to the relevant colleagues and also agreed to send the qualitative questionnaire (potentially with the delay).

- o Italy - we are organizing a meeting during the week starting from 10/07 with the Italian IM in order to gather the relevant data.

- o Portugal - IM asked for an editable version of the questionnaire and requested 6 month for filling out the quantitative questionnaire. We replied that it is not possible and kindly requested their participation in the agreed timeframe. No further feedback has been received from IM side.

- o Spain - we have contacted the IM several times by email and by phone. We established contact by phone on 20th of June, the individual stated that they do not have enough time in their department for providing such an amount of information within the deadline. He requested to send to him another email (on the 20th) in order to assess the possibility to provide the information requested. Up to date he has not provided any feedback to our gentle reminder and between 26th June until 04th July we tried to contact him several times by phone without success.

- o Sweden - asked a possibility to propose us a new deadline based on their time schedule. We are waiting for their feedback.

- o UK - A message was dropped in voice mail (in order to obtain news from them). They emailed us saying that they will provide us with the data before the deadline.

- Until 04 of July 2017 we have received total of 29 responses to the qualitative survey:

  - o 6 MA from LV, PT, SI, CZ, RO and IT

  - o 13 IM from IT, GR, HU (NIF Zrt.), CZ, BG, DK, AT, DK, HU (MÁV), SK, PT, SE, HR

  - o 10 MoT from EE, DE, BE, BG, HR, LT, ES, LV, CZ, LT

In the same period we have received only 2 responses, from Slovenia and Fermen A/SMulticountry DE/DK stakeholder - Fehmarnbelt tunnel), to the quantitative questionnaire.

- Even though the deadline for qualitative survey has passed (30 June 2017), based on the undertaken discussion with the stakeholders and their requests for the extension of deadline, we are still expecting to receive additional replies throughout month of July.

- To monitor the status of the stakeholders consultation a monitoring sheet was prepared. The second monitoring sheet is attached to this report.

Data collection - DG REGIO and INEA database

- In the course of the data mining and disaggregation process, we have identified several projects with missing documentation and accordingly additional data collection activity has been organised with the DG REGIO and INEA.

- In reporting period, DG REGIO provided us with the additional project documentation for 35 projects.

- From INEA we have requested additional documents for 47 projects.

- The response was very fast and effective what enabled us to continue the data mining process uninterruptedly.

- In total, at this moment we are still missing information for 7 DG REGIO projects.

  - For the 2 Croatian projects (Zapresic-Zabok rail reconstruction and upgrade and Railway line Vinkovci - Vukovar) it has not been possible to collect information from DG REGIO server as we do not have the project ID. We have requested information on this projects from the stakeholders

  - For the remaining projects (2 FR -2 DE-1 IT) no information was available in the project folders.

Data manipulation (DG REGIO and INEA) and compilation of the database

In the reporting period the team continued the DG REGIO and INEA data manipulation process by applying following activities:

- **Data mining and disaggregation** was finalized for 379 projects (70% of the total 544 projects), out of which 262 from DG REGIO and 117 from INEA database. It included reading all the available documents (more than 7.500 documents) and extracting the required data into the working database.

- **Data cleaning** - the team has performed a detail revision of all extracted data and removed all invalid data points from a dataset.

- **Data organization** - the team has performed classification (in 3 Tiers) and organization of data sets to make them more useful for the planned analysis to be undertaken in order to evaluate the unit cost in rail infrastructure projects.

- **Data integration and storage** - the team has performed integration of all working databases and storing the final data into the database. We have merged received quantitative data from the Stakeholders (Slovenia and Fermen) into the final database.

Summary:

- **DG REGIO database -** data extraction (262 projects, 80% of the total DG REGIO) was more complicated then INEA data, as almost no data was presented in aggregated format. All data was available only in paper document format, while the number and type of documents varied among projects. On average for each project we have manipulated 20 documents.

- All projects from the INEA database for TEN-T 2007 – 2013 programme were manipulated (117 projects). INEA database required lower amount of time as the data was presented in an aggregated format (excel). **The level of data availability from TEN-T 2007 – 2013 programme was limited to Tier 1.**

- Until now we haven't processed any of the total 96 projects from the CEF 2014 programme. We are still missing the projects there were subsequently sent to us by INEA representative on 29 June 2017.

| 2. | **Issues and challenges** |

- In case if the data received from Stakeholders will be different from the one gathered from DG REGIO/ INEA database, we have decided to use the data from stakeholders as a relevant one, while at the same time making additional checks in order to determine the source of difference (e.g. different level of reported cost estimation)

- Stakeholders' commitment to the process is of utmost importance and PwC is putting maximum commitment in order to receive as high as possible response rate and data coverage.

  - The collection of qualitative data from stakeholders has been progressing slowly.

The total response rate is at 24%. The highest rate of response is recorded from IM (34%), followed by MoT (23%) and MA (17%).We plan to include only preliminary analysis in the first progress report. In terms of geographical coverage, currently the responses from IM account 46% of the MSs , from MoT 38% and from MA with 23%.

- The total expected rate of response for the qualitative questionnaire based on the commitments form the stakeholders is 40%. Per stakeholders we are expecting following rates - 56% for IM (including responses from BE, FI, DE, LT, LV, UK, TP Ferro), 34% for MoT (including responses from DK, FI, GR, HU, NL) and 30% for MA (including responses from HR,GR,SK,HU,PL). It would result in a geographical coverage of 73% for IM, 58% for MoT and 42% for MA.

- For the collection of quantitative data, stakeholders already started asking for extension of the deadline (deadline is July 14th). For the time being, we are still not able to plan when we will get full set of information.

- EC has been timely informed on the difficulties in qualitative and quantitative data collection from the countries (IM) with the highest number of projects (Poland and Spain) and necessary steps have been taken. EC has engaged Regional Units in the data collection process and the requests have been submitted from their side to the Stakeholders.

Issues relevant for the production of the First progress report

- **The Tier 2 and Tier 3 data availability is limited (based on DG REGIO and INEA database).** In case if the rate of response from the Stakeholders will not be up to the required level, it is necessary to define how to proceed with the analysis.

  - The Tier 2 data availability depends on the characteristics of the information, even if, it remains low (based on DG REGIO and INEA database). For Tier 2, 40 % of projects are with partially available data, while 60% have no data available. However, the trend differs significantly analysing technical information (60% of projects contain these information) or economic ones (only 15%).

  - The Tier 3 data availability is limited (based on DG REGIO and INEA database). Regarding technical information, a half of projects are with partially available data, while from the economic side, 6% of projects present information. Generally, 29% projects have partially available data and 71% no data available.

- It is necessary to agree how to make the geographical analysis for the **countries without projects** on the project list (Ireland and Estonia) and the ones for which we do not have infrastructure projects in the list (e.g. Denmark has only ERTMS projects)

  - The potential mitigation measures are either to apply the similarity rule and use the analysis from the countries with similar features, or to make an additional detail research through the meetings with the Infrastructure Managers and other relevant Authorities and prepare the Case Studies.

- Although **the methodology** initially agreed did not undergo additional revision during the inception phase, we were planning to fine-tune the analysis and development phase approach at the end of the data collection phase and to present it at the first progress report. However, due to the higher number of projects then initially planned under the ToR and the need to screen all papers to extract data and to involve a large number of stakeholders, the data collection phase as a whole is taking more time than planned.

	o Accordingly, since the data collection phase was not fully completed up to the time of the First Progress report production (the data received/ manipulated until 04 of July 2017 is analysed in the report), **the fine-tuned methodology is not part of this report.** o If needed, we will produce a separate report in September that will present the fine-tuned methodology. o We are verifying the availability of our senior statics exeprt (Diego Artuso) to join the Steering Group meeting to start discussing the statistical analysis. I ■ We are preparing the First progress report that will be consisted of following parts: 　o **Chapter 1** provides a general overview of the project 　o **Chapter 2** describes the process and the results including the elaboration of the aim, approach, status and results of preliminary analysis from the qualitative survey. 　o **Chapter 3** elaborates the data collection and manipulation process. In includes assessment of the data mining, cleaning, organisation progress and data processing issues. 　o **Chapter 4** provides a description of cost sub-categories, presentation of lower, median and upper values and list of potential issues to be mitigated in the stage of data collation, calibration and reaggregation. 　o **Chapter 5** presents the draft list of the potential Case Studies to be developed in the next period and the reasoning behind the proposal (e.g. reduced data availability for Tier 3 tunnels and bridges, lack of information for particular countries, etc.).We will also include the description of how we are planning to take stock of existing documents relevant for the study (e.g. WP10 final report, Study on ERTMS, etc) 　o **Chapter 6** presents next steps.
3.	**Next steps**
	• We are continuing stakeholders consultation process. Follow up calls and e-mails will continuously be sent to ensure participation from all 3 categories of the stakeholders with the highest emphasis on the IM. • The First Progress report will be delivered by 20th of July 2017. • The third Steering Group meeting will be organised soon after the delivery of the report. • In case of a need, a bilateral meeting with the stakeholders will be organised. • Considering the level of data availability that is required for the production of the geographical fact sheet, we would like to postpone its delivery from September to October. • As soon as we have compiled the database and have a clear overview of the data available, we will organise a meeting with the JASPER in order to discuss possible methodologies.

Addressees	**Kyriacos Ktenas (DG REGIO – F.1)**
Author	**PwC**
Project Name	**Assessment of unit costs (standard prices) of rail projects (Capital Expenses)**
Subject	**PROGRESS NOTE 07 August 2017**

1. **State of play and key progress from the last Progress Note (July 7th) to August 3rd**

Data collection - Stakeholders consultations

- The status of the stakeholders consultation process is synthetized in the following table

MS	STK	Qualitative	Quantitative	Commitment	Comment	Performance
AT	IM	Yes	No	Yes	In progress	Medium-Low
	MA	No		No	No contact was established	Low
	MoT	No		No	No contact was established	Low
BE	IM	Yes	Partial (38%)	Yes	In progress	Medium-High
	MoT	Yes		Yes		High
BG	IM	Yes	Yes (100%)	Yes	To be revised	High
	MA	MoT is the MA		Yes		High
	MoT	Yes		Yes		High
HR	IM	Yes	Partial (60%)	Yes	In progress	Medium-High
	MA	MoT is the MA		Yes		High
	MoT	Yes		Yes		High
CZ	IM	Yes	No	Yes	In progress	Medium-Low
	MA	Yes		Yes		High
	MoT	Yes		Yes		High
DK	IM	Yes	No	Yes		Medium-Low
	MoT	No		Yes	In progress	Medium-High
EE	IM	No	No	Yes	In progress	Low
	MA	Yes		Yes		High
	MoT	No		No	No contact was established	Low
FI	IM	No	Partial (50%)	Yes	In progress	Medium-Low
	MoT	No		Yes	In progress	Medium-High
FR	IM	No	No	No	They're not responsive	Low
	MoT	No		No	They're not responsive	Low
DE	IM	No	No	No	To be confirmed	Medium-Low
	MoT	No		Yes	In progress	Medium-High
EL	IM	Yes	No	Yes	In progress	Medium-Low
	MA	No		No	No contact was established	Low
	MoT	No		No	In progress	Medium-Low
HU	IM-MAV	No	No	Yes	In progress	Low
	IM-Gysev	Yes	Yes (100%)	Yes		High
	MA	No		No	No contact was established	Low
	MoT	No		No	No contact was established	Low
IT	IM	No	No	Yes	In progress	Medium-Low
	MA	No		No	No contact was established	Low
	MoT	No		No	No contact was established	Low
IE	IM	No	No	Yes	Suggested asking for more recent projects	Medium-Low
	MoT	Yes		Yes		High
LV	IM	No	Yes (100%)	Yes	Only Tier 1 info was included, PwC Latvia has been requested to intervene	High
	MA	Yes		Yes		High
	MoT	Yes		Yes		High
LT	IM	No	No	Yes	PwC Lithuania will intervene	Medium-High
	MA	Yes		Yes		High
	MoT	Yes		Yes		High
LU	IM	No	No	Yes	In progress	Medium-Low
	MoT	No		No	In progress	Medium-Low
NL	IM	No	No	No	No contact was established	Low
	MoT	Yes		Yes		High

PL	IM	No	No	No	Waiting for the STK confirmation to provide a reduced number of projects (07-13)	Low
	MA	No		No	Looking for the right contact	Low
	MoT	No		No	In progress	Medium-Low
PT	IM	Yes	Partial (40%)	Yes	Only information on the period 2007-2013 is available	Medium-High
	MA	No		No	No contact was established	Low
	MoT	Yes		Yes		High
RO	IM	Yes	No	Yes	In progress	Medium-Low
	MoT	No		No	No contact was established	Low
ES	IM	No	No	No	Too many projects to be requested, no cooperation /response from the IM	Low
	MA	Yes		Yes		High
	MoT	No		No	No contact was established	Low
SK	IM	Yes	No	Yes	In progress	Medium-Low
	MA	Yes		Yes		High
	MoT	Yes		Yes		High
SI	IM	No	Partial (69%)	Yes	In progress	Medium-High
	MA	No		No	No contact was established	Low
	MoT	No		Yes	In progress	Medium-High
SE	IM	Yes	Partial	Yes	In progress	Medium-High
	MA	No		No	No contact was established	Low
	MoT	No		No	MoT responded to the IM survey	Medium-High
UK	IM-Translink	No	Yes (100%)	Yes	Only Tier 1 info was included	Medium-High
	IM-Network Rail	No	No	No	Still in need to reach the right contact within Network Rail	Low
	MoT	No		No	No contact was established	Low
IE/UK	IM-DRDNI	-	Via Translink (100%)	Yes	The info was gathered through Translink	N/A
IT/FR	IM-TELT	-	No	Yes	In progress	Medium-Low
FR/ES	IM-TP Ferro	-	No	Yes	In progress	Medium-Low
FR/ES	IM-GEIE	-	No	Yes	In progress	Medium-Low
IT/AT	IM-BBT	-	No	Yes	In progress	Medium-Low
EE/LT/ LV	IM- RB	-	No	No	No contact was established	Low
DE/DK	IM-Femern	-	Yes (100%)	Yes		High

- *Highlighted in red those infrastructure managers for which the intervention of the EC might be necessary.*

Lithuanian and Latvian railways required support from PwC local offices to extract project information from the documents they have at their premises.

We met Lithanian Railways (LR). Summarising the outcome of the meeting:

- Data availability and its format will be problematic. For older projects before 2010 (there were 5 such identified), all information is stored in archives in the paper format and some of it in different city. For some of these projects which are dated 2002 and 2004 LR personnel was doubtful if we will even be able to find all the documentation required.

We will start supporting them and we will report on the progress.

Data collection and manipulation (DG REGIO and INEA) and compilation of the database

Out of total 544 projects, 403 have already been analysed in detail, to extract the data necessary for the assignment.

Before proceeding with the analysis, the data was adjusted to ensure that the calculations were performed on a uniform and homogenous basis:

- **Data normalisation**, in order to express them in a fictional currency representing the (weighted) average EU28 in terms of purchasing power parity;

- **Harmonisation of technical data,** project length was adjusted so that the difference in single-track and double-track investments is taken into consideration;

- **Segmentation of data,** based on the main type of works envisaged, projects were classified into eight categories (i.e. signalling, hub and large structural projects, electrification, new lines not impacted by complexity, new lines affected by complex civil construction works, rehabilitation and upgrading of existing lines not impacted by complexity, rehabilitation and upgrading of existing lines based on complex civil construction works, not specified) in order to compare costs related to similar investments

Testing methodological approaches

The project-related information have been normalised and harmonised for preliminary tests to be performed to:

- understand if the information currently available and the current level of detail are sufficient to identify a clear and significant correlation between infrastructure length and costs, per project type;

- spot the presence of observations out of scale and understand the reason behind their unexpected values;

- define and, to the extent possible, confirm methodological steps that enable to achieve the desired outcomes by statistically analysing project-level data;

- identify the limits of the analyses and, coherently, the steps to overcome them.

The approach consists in the segmentation of projects per project-type and analysing similar projects at cluster level. The approach has the benefit of being applicable to a dataset with relatively low detail of data. Nonetheless it is less and less reliable the wider the category of projects analysed.

Summary:
The preliminary results are promising and validate the approach. Indeed, although a number of outliers were identified at different project category level, the main reasons for their presence have been detected. Coherently, the analysis of the dispersion of project costs per category type shows that a reasonable, average cost range can be identified for the majority of categories. For those for which it has not been possible to derive a statistical relation, further details are required. The analysis shall indeed be performed on cost items.

2.	**Third SGUCR meeting**

The meeting was held on 26th July 2017 in Brussels, focusing on the First Progress Report and the related activities performed to date.

The preliminary results of the data collection and analysis were presented, both relative to the qualitative and quantitative analyses.

Data availability was confirmed to be the key aspect on which to focus the effort:

- The availability of data on projects from DG REGIO and INEA documentation shall be complemented with stakeholders' inputs. Would this not be the case, the analyses could only be based on Tier 1 data.

- Similarly for the qualitative analyses, the response rate has been overall significant, but the involvement of stakeholders has to be ensured in the following months, to reach a sufficient data coverage.

The European Commission will increase its effort to convince all relevant parties to respond to the request for data and information. At the same time, PwC continues to ask stakeholders to take an active part to the exercise, as well. Attached, the responsiveness from each MS is reported, together with the need for actions to be taken to stimulate participation.

After the presentation of the data collection process, the preliminary results of the qualitative analysis were shown and commented. In particular, it is suggested to consider identifying some measures to test the responses of the stakeholders.

Similarly, the methodologies tested on the quantitative data were presented and discussed. The approach to normalisation of costs is suggested to be compared with different methodologies. The initial analyses provided promising results; the relative methodology is to be further discussed and validated with the SGURC.

3.	**Next steps**

- We are continuing stakeholders consultation process. Follow up calls and e-mails will continuously be sent to ensure participation from all 3 categories of the stakeholders with the highest emphasis on the IM.

- In case of a need, a bilateral meeting with the stakeholders will be organised.

pwc		European Commission

Addressees	**Kyriacos Ktenas (DG REGIO – F.1)**
Author	**PwC**
Project Name	**Assessment of unit costs (standard prices) of rail projects (Capital Expenses)**
Subject	**PROGRESS NOTE 06 September 2017**

1.	**State of play and key progress from the last Progress Note** (August 3rd) **to September 6**th

Data collection - Stakeholders consultation

All stakeholders were contacted to screen the status of the data provision process and to offer assistance, when necessary. In most of the cases, it was possible to contact the interested party and ensure cooperation from the stakeholder to receive the necessary information in a relatively short time. In a few cases, the activities were hampered by stakeholders' absence due the summer vacations.

As a number of stakeholders indicated the number of projects to be too burdensome, to have more timely responses and ensure a greater cooperation, priority was given to the projects in the 2007-2013 programming period. Stakeholders often claimed it was not possible to deliver the Tier3 level of detail due to different data organisation issues and to time/resource constraints. Whenever possible, we offered to support with our local teams. When not feasible to reach Tier3 level, stakeholders were asked to limit the data provision at Tier2 level.

Hereunder, the main cases where some peculiarities occurred in the collaboration process are described:

Austrian IM: although had initially agreed on participating in the study, the collaboration has not been given the expected result yet. The IM had issues in gathering the information and sent a limited number of information per each project. They are checking if they can provide additional information.

Czech IM: the IM provided us with a very detailed database comprising all of the projects requested, in Czech. The local PwC office was involved in order to analyse such projects and transfer the raw data to the database. Focus was given to the projects completed in the 2007-2013 programming period, as more complete information on both estimated and final costs is available.

French and German IM: in these countries the IMs were reached but no confirmation was provided as to whether the IM would participate in the study. In order to ensure full cooperation a compromise was reached and priority in the data collection was given to the projects in the 2007-2013 programming period and Tier 1 and 2 where Tier 3 was too difficult to be retrieved. More time will be needed for these IM to retrieve the data, and they say they will try to send the quantitative data in the coming weeks.

Greece: Greek DG REGIO projects were analysed from the local PwC team, as they were available in Greek. The Greek IM has been responsive and cooperative and requested the support of PwC to extract the data from their databases. Our local team is currently coordinating with the IM. The IM asked for an extension of the deadline due to the holiday period. They will be working together with our Greek colleagues starting from later this week.

Poland: the Polish IM has been responsive and cooperating, however Poland has a very high number of projects to be requested. Hence, focus has been given to the project carried out in the

2007-2013 programming period, and Tier 1 and Tier 2 where Tier 3 is too difficult to be reported. A deadline extension was agreed to gather all the needed information. For what concern the Polish Managing Authorities it was impossible to contact of a person that could answer to the questionnaire.

British IMs: two different IMs have been contacted: The Translink (responsible for North Ireland) and Netowork Rail. The IM of Translink gave information only on Tier 1, as it was not possible to include further details. For what concern Networkrail, it was inpossible to establish the person in charge and able to fill in the information. Therefore, it is not clear if we will received the information.

Latvian and Lithuanian IMs: in order to extract the information necessary, the support from local PwC offices was necessary. PwC teams had cooperated with the IMs to extract project information, available in the local language, and subsequently elaborated such data. The consultation is currently on going.

Spanish IM: the IM has not been very keen to participate in the study, even after offers from PwC to reduce the amount of detail required or number of projects to be collected, providing only a generic report, which contains information on unit costs. The document can be used anyway to support the study, but we are anyway trying to find a compromise to have ADIF actively participating.

The status of the stakeholders consultation process is synthetized in the table attached.

<u>Data manipulation and compilation of the database</u>

Once the quantitative data was received from the different stakeholders, such data was revised, in order to check the quality of the information and whether inconsistencies were present across tiers. If that were the case, the data was sent back to the stakeholders to correct and provide further clarifications. Once the revised information was received, it was transposed from the stakeholders' questionnaire in the database to be used for the analyses.

Before starting the analysis, it is necessary to have unique values for each project. Therefore a comparions of the information collected from the two sources, DG REGIO / INEA and from the stakeholders is being performed. This control ensures the triangiulation of the informarion and a higher reliability of the data. In case the difference across data sources is meaningful, the stakeholders are contacted to provide explanations or corrections where needed. When the difference is not significant, priority is given to the data retrieved from the stakeholders, as it is more detailed.

Normalisation

Following the suggestions from the First Progress Meeting, it was researched whether the application of the HCPI would have resulted in unbiased results, to homogeneise the differences across MSs in Europe. Being the HCPI a consumer-based index, it does not include the construction sector, as it only refers to consumer-related sectors. As a result, its application would not fit the purpose of the study, as it necessarily excludes the sector we aim to notmalise. The Purchasing Power Parity is more general, but thanks to this cross-sectorial nature better fits the purpose.

2.	**Next steps**
	• In the coming weeks, the main task to be carried out remains the collection of the quantitative data. In particular as a number of stakeholders required more days than expected to provide us with the necessary information. Therefore, the stakeholders' consultation process will continue. In order to be able to reach all the stakeholders, follow up calls and e-mails are constantly performed. • The ultimate deadline for the provision of data and information from stakeholders is yet to be decided. While on one side a longer period would ensure a higher response rate, analyses have to be made on a stable and finalised database. • When the deadline for the provision of information is defined, a meeting with the Commission and the Steering Group will be called to present the database structure and present the proposed methodology, tailored on the data available then.

MS	STK	Qualitative	Quantitative	Commitment	Comment	Performance
AT	IM	Yes	Yes		Sent one project, which is under revision process. They will check if they can send more	Medium-Low
	MoT	Yes		Yes		High
BE	IM	Yes	Yes	Yes	They sent 4 projects. The will provid the rest by mid-September	Medium-High
	MoT	Yes				High
BG	IM	Yes	Yes	Yes	They sent all the projects.	High
	MA	MoT is the MA				High
	MoT	Yes				High
HR	IM	Yes	Yes	Yes	The sent 3 projects, which are under revision process. They will check if they can send more projects	Medium-High
	MA	MoT is the MA				High
	MoT	Yes				High
CZ	IM	Yes	Yes	Yes	They sent all the infomration in Czech. PwC CZ provided the analysis of the 07-13 projects	High
	MA	Yes				High
	MoT	Yes				High
DK	IM	Yes	Yes	Yes	One project was sent, which is under revision process. The other projects were not sent because of lack of unit cost data, we asked if they had other projects available	Medium-High
	MoT	No		Yes		Medium-Low
EE	IM	Yes	Expected	Yes	Seemed cooperative at first, not responsive for some time.	Medium-Low
	MA	Partial		Yes		Medium-High
	MoT	No				Medium-Low
FI	IM	No	Yes	Yes	They sent 3 projects that need to be revised. Unable to provide information on others.	Medium-High
	MoT	No		Yes		Medium-High
FR	IM	No	Expected	Yes	Have only recently appointed the person in charge of the task. Will try to extract data up to Tier2.	Medium-Low
	MoT	No		No		Medium-Low
DE	IM	No	Expected	Yes	Have only recently appointed the person in charge of the task. Will try to extract data up to Tier2	Medium-Low
	MoT	Yes				Medium-High
EL	IM	Yes	Expected	Yes	They are cooperating with our local team to get the detailed projects ready in the coming weeks.	Medium-High
	MA	Yes				High
	MoT	No				Medium-Low
HU	IM-MAV	No		Yes	They got us in touch with NIF, which is responsible for the implementation of the Projects of the 2007-2013 period.	Medium-Low
	IM-NIF	No			We contacted them this week and we are waiting for a reply	-
	IM-Gysev	Yes	Yes		They sent one project , which needs to be revised	Medium-High

MS	STK	Qualitative	Quantitative	Commitment	Comment	Performance
	MA	No				Medium-Low
	MoT	No				Medium-Low
IT	IM	No	Expected	Yes	Recently came back from annual leave. Activity in progress.	Medium-Low
	MA	Yes				High
	MoT	Yes				High
IE	IM	Yes	Yes			High
	MoT	Yes				High
LV	IM	Yes	Yes	Yes	The sent project only included Tier 1 info, PwC Latvia is working with the Latvian IM to retrieve info on Tier 2 and 3	Medium-High
	MA	Yes				High
	MoT	Yes				High
LT	IM	Yes	Yes	Yes		High
	MA	Yes				High
	MoT	Yes				High
LU	IM	No	Expected	Yes	Seemed cooperative at first, not responsive for some time.	Medium-Low
	MoT	Yes				High
NL	IM	No	Expected	Yes	In contact with them. Sent little information, which we are trying to have integrated.	Medium-Low
	MoT	Yes				High
PL	IM	No	Expected	Yes	Asked us to revise our requests and decrease the number of projects requested. We agreed on the projects in the 2007-2013 period, and if not possible only info on Tier 1 and Tier 2	Medium-Low
	MA	No			No answer from their side	Low
	MoT	No				Medium-Low
PT	IM	Yes	Yes	Yes	Only information on the period 2007-2013 is available	Medium-High
	MA	Yes			The MA is not in charge of the planning of rail projects, we are in contact with another department	Medium-High
	MoT	Yes			The MoT is not in charge of the planning of rail projects	High
RO	IM	Yes	Expected		Should provide the projects starting from next week.	Medium-High
	MA	Yes				High
	MoT	No				Medium-Low
ES	IM	No	No	No	Although we offered support to reduce the amount of work, they are not likely to provide any data on projects. Trying to involve them through out local team.	Low
	MA	Yes				High
	MoT	No				Low
SK	IM	Yes	In progress	Yes	They will send the raw data of the projects by this week. PwC Slovakia will collaborate with us to extract the data	Medium-High

MS	STK	Qualitative	Quantitative	Commitment	Comment	Performance
	MA	Yes				High
	MoT	Yes				High
SI	IM	Yes	Yes	Yes	They sent 10 projects. The will not be able to send more projects before the end of the year	High
	MA	Yes				High
	MoT	Yes		Yes		High
SE	IM	Yes	Yes	Yes	Only one project is missing, which will be delivered soon	High
	MoT	Yes			IM responded also for MoT	High
UK	IM-Translink	-	Yes		Only Tier 1 info was included, no possibility to include further detail	Medium-Low
	IM- Network Rail	No			No answer from their side	Low
	MoT	No				Medium-Low
IE/UK	IM-DRDNI	-	Translink		The info was gathered through Translink	N/A
IT/FR	IM-TELT	Yes	Yes	Yes		High
FR/ES	IM-TP Ferro	-	Yes	Yes	One project was sent, which is under revision process	Medium-High
FR/ES	IM-GEIE	-	Expected	Yes		Medium-Low
IT/AT	IM-BBT	Yes	Yes	Yes		High
EE/LT/LV	IM- RB Rail AS	-				Medium-Low
DE/DK	IM-Femern	Yes	Yes	Yes	They sent one project, which is under revision process	Medium-High

Addressees	**Kyriacos Ktenas (DG REGIO – F.1)**
Author	**PwC**
Project Name	**Assessment of unit costs (standard prices) of rail projects (Capital Expenses)**
Subject	**PROGRESS NOTE 29 September 2017**

1.	**State of play and key progress from the last Progress Note (September 7th) to September 29th**

First progress report revision

The Commision provided the comments on the draft of the First Progress Report on the 19th September. The Report was updated accordingly and sent back to the Commission for revision with a summary of the changes made on the 25th September. Specifically, the updated version of the report includes additional information on:
- The backlog of the data processing activity and related collection of data and information;
- The analysis that will be performed on the data gathered;
- The role of scientific research and the ancillary data sourcing within the study.

Data collection - Stakeholders consultation

A significant progress has been reported on the data collection. The process was completed for a consistent amount of stakeholders, compared to the previous period: Belgium, Czech Republic, Estonia, Finland, Latvia, Lithuania, and Netherlands.

For time issues, the consultation has been focused on the stakeholders that showed higher involvement in the project. All of them were recontacted in order to have updates and to offer assistance, in case of necessity. In most of the cases, they renewed their intention to take part in the study and provide us with the data in a short time, in accordance with their workload.

Hereunder, the main cases where some peculiarities occurred in the collaboration process are described:

Austrian IM: A more strict collaboration with the IM has been established. The IM provided the revision of the project already submitted and ensured that the data on three additional projects will be provided compatibly with the information available. However, they were not able to estimate a deadline for the submission of the information.

French IM: the IM was responsive, but, as in the previous period, no confirmation was provided as to whether the IM would participate in the study. A conference call to have updates on the data collection process has been organised for next week.

German IM: the IM was cooperating and provided a huge amount of data in excel format. However, the information does not fit the cost breakdown structure for the data collection process. They are currently checking if they can provide additional information.

Greece: Although the PwC local team offered assistance to extract the data from the databases, the IM was not able to provide information more detailed than the one reported in the Application Form. The Greek Ministry of Transport is coordinating internally in order to answer to the questionnaire.

Poland: the Polish IM was responsive and confirmed its collaboration on the data collection process. It was not possible to estimate the new deadline for the submission of the projects, as the

data gathering has started recently. We will keep in touch for updates in the following week.

Romania: Although the Romanian IM seemed cooperative, they did not provided the data by the agreed deadline and they are no more responsive. PwC will continue to try to contact them in the following week.

United Kingdom: It has not been possible to identify the right person to be involved in the consultation within Network Rail. A new contact has been established by the Commission, within the Office for Rail and Road (ORR) in the UK. They have been cooperative and will check if they can provide some information. The British Ministry of Transport filled out the qualitative questionnaire.

The country detail of the stakeholders consultation process is synthetized in the table attached.

Compilation of the database, data manipulation and analysis

The data received from the different stakeholders was revised to check the quality of the information and whether inconsistencies were present across tiers. The inconsistencies have been clarified for most of the projects concerned and when possible, a higher level of detail has been included, thanks to the collaboration of the IMs. As in the previous period, the information received was transposed from the stakeholders' questionnaires to the global database and compared with the data collected from the EC (e.g. project application and grant agreements). This allowed both to test the reliability of the data filled out manually in the database and to double-check the correspondence with the projects provided by the stakeholders.

A unique database has been filled out both with the data received from the stakeholders and the information collected from the two sources, DG REGIO and INEA. The data cleaning and manipulation aimed to obtain comparable data. The unit costs and the percent incidence of the costs components on the total cost have been calculated for the three tiers.

A preliminary data analysis was performed both at European and country level, for the different line categories (high speed or conventional) and type of work (new line construction, rehabilitation and upgrading, signalling or electrification). The analysis aims to test the approaches for a unit cost range for the same categories within the same cluster. Projects showing values considerably bigger or smaller than the range expected, have been looked in detail, aiming to identify whether the reason for the discrepancies or a possible error in the data filling.

This laid the basis for the identification of the main costs drivers and eventually, for a multivariate linear regression analysis. The multivariate regression analysis is being tested against the database – which is nonetheless being updated with late inputs – to test the possibility to derive a generally applicable rule that links the infrastructure features to its cost.

Final report

A first draft of the structure of the final report has been elaborated. The same structure is suggested to be used to draft the second progress report – due by the end of October. The difference will be in the level of completion of the information included in the paragraphs. The structure is attached for approval.

2.	Next steps
	- In the coming weeks, the main focus will be given to carry forward the multicriteria analysis and the multivariate analysis (top-down approach) as well as to proceed through bottom-up statistical approaches (frequency based on smaller clusters). - The collection of the quantitative data will continue in the following weeks, in order to collect the data on the missing countries. In particular for the stakeholders which required more days than expected to provide us with the necessary information. In order to be able to reach all the stakeholders, follow up calls and e-mails are constantly performed. - A meeting is suggested to be scheduled on 18th October, to discuss the progress of the study, to present the structure of the database, the data quality and population as well as to anticipate the content of the second progress report.

MS	STK	Qualitative	Quantitative Response	Commitment	Comment	Performance
AT	IM	Yes	Yes		They sent one project. They will send three more projects.	Medium-Low
	MoT	Yes				High
BE	IM	Yes	Yes		They sent all the projects.	High
	MoT	Yes				High
BG	IM	Yes	Yes	Yes	They sent all the projects.	High
	MA	MoT is the MA				High
	MoT	Yes				High
HR	IM	Yes	Yes	Yes	The sent 3 projects. They will check if they can send more projects	Medium-High
	MA	MoT is the MA				High
	MoT	Yes				High
CZ	IM	Yes	Yes		They sent all the infomration in Czech. PwC CZ provided the analysis of the 07-13 projects	High
	MA	Yes				High
	MoT	Yes				High
DK	IM	Yes	Yes		One project was sent, its information needs to be revised. The other projects were not sent because of lack of unit cost data, we asked if they had other projects available	Medium-Low
	MoT	Yes			The IM will fill it out	High

23

MS	STK	Qualitative	Quantitative Response	Commitment	Comment	Performance
EE	IM	Yes	Yes			High
	MA	In progress		Yes	Questionnaire to be completed	Medium-High
	MoT	No				Medium-Low
FI	IM	Yes	Yes			Medium-Low
	MoT	Yes			The IM and MoT coordinated to fill out the questionnaire.	Medium-High
FR	IM	No	Expected	Yes	In progress – A conference call is organised next week to update on data retrieve process	Medium-Low
	MoT	No			The IM will fill it out	Medium-Low
DE	IM	No	No	Yes	They provided some data which cannot be used in our analysis, they will check if they can provide more data.	Medium-Low
	MoT	Yes				Medium-High
EL	IM	Yes			They will send the projects in the coming weeks	Medium-Low
	MA	Yes			The MA is not in charge of the planning of rail project	High
	MoT	No	Expected	Yes	In progress	Medium-Low

MS	STK	Qualitative	Quantitative Response	Commitment	Comment	Performance
HU	IM-MAV	No		Yes	Have been cooperative, but have no information on projects, therefore put us in contact with NIF.	Medium-Low
	IM-NIF	Yes	Expected	Yes	Contact provided by MAV, they will work to provide the data on three projects	Medium-High
	IM-Gysev	Yes	Yes			High
	MA	Yes				Medium-Low
	MoT	Yes			The Agency is not in charge of the planning of rail project	Medium-Low
IT	IM	No	Expected	Yes	Some information have been included, some is expected to be retrieved soon.	Medium-Low
	MA	Yes			The MA is not in charge of the planning of rail projects	High
	MoT	Yes			The MoT is not in charge of the planning of rail projects	High
IE	IM	Yes	Yes			High
	MoT	Yes				High
LV	IM	Yes	Yes		The sent project only included Tier 1 info, PwC Latvia is working with the Latvian IM to retrieve info on Tier 2 and 3. Should send initial info next week.	High
	MA	Yes				High
	MoT	Yes				High
LT	IM	Yes	Yes			Medium-High

MS	STK	Qualitative	Quantitative Response	Commitment	Comment	Performance
	MA	Yes				High
	MoT	Yes				High
LU	IM	No			Not responsive	Medium-Low
	MoT	Yes				High
NL	IM	Yes	Yes			High
	MoT	Yes				High
PL	IM	No	Expected	Yes	They are working on the data collection.	Medium-Low
	MA	No			They asked to resend the email with the information on the project	Medium-Low
	MoT	Yes				Medium-Low
PT	IM	Yes	Yes	Yes		Medium-High
	MA	Yes			The MA is not in charge of the planning of rail projects, we are in contact with another department	Medium-High
	MoT	Yes			The MoT is not in charge of the planning of rail projects	High
RO	IM	Yes	Expected		In progress	Medium-Low
	MA	Yes				High

MS	STK	Qualitative	Quantitative Response	Commitment	Comment	Performance
	MoT	No			In progress	Medium-Low
ES	IM	No	No	No	Although we offered support to reduce the amount of work, they are not likely to provide any data on projects – we have received alternative documentation we can use to integrate, though.	Low
	MA	Yes				High
	MoT	Yes			The MoT is not in charge of the planning of rail projects	High
SK	IM	Yes	In progress	Yes	The IM is collaborating with PwC SK to fill out the excel files. The process is almost completed	High
	MA	Yes				High
	MoT	Yes				High
SI	IM	Yes	Yes	Yes		High
	MA	Yes			The MA is not in charge of the planning of rail projects	High
	MoT	Yes		Yes		High
SE	IM	Yes	Yes	Yes	Only one project is missing, which will be delivered soon	High
	MoT	Yes			IM responded also for MoT	High
UK	IM-Translink	No	Yes		Only Tier 1 info was included, no possibility to include further detail	Medium-High
	ORR				Contact established by the Commission	N/A
	IM-Network Rail	No			Unable to establish the right contact	Low

MS	STK	Qualitative	Quantitative Response	Commitment	Comment	Performance
	MoT	Yes				High
IE/UK	IM-DRDNI	-	Translink		The info was gathered through Translink	N/A
IT/FR	IM-TELT	Yes	Yes	Yes		High
FR/ES	IM-TP Ferro	-	Yes			Medium-High
FR/ES	IM-GEIE	-		No		Medium-Low
IT/AT	IM-BBT	Yes	Yes	Yes		High
EE/LT/LV	IM- RB Rail AS	-			In progress	Medium-Low
DE/DK	IM-Femern	Yes	Yes	Yes	They sent one project, which is currently under revision	Medium-High

Addressees	**Kyriacos Ktenas (DG REGIO – F.1)**
Author	**PwC**
Project Name	**Assessment of unit costs (standard prices) of rail projects (Capital Expenses)**
Subject	**PROGRESS NOTE 27 October 2017**

1.	**State of play and key progress from the last Progress Note (September 29th) to October 27th**

Data collection - Stakeholders consultation

As in the previous period, the consultation focused on the main stakeholders involved in the projects. In the process, the infrastructures managers were recontacted. As a result, the **Austrian Infrastructure Manager** sent the required information. Nevertheless, in most of the cases where data collection is still pending (or partially pending) more time is needed to provide us with the quantitative information requested.

In particular:

French IM: the IM was responsive, and agreed on participating in the study. They started with a first data screening and provided us with a list of the projects, which they are going to analyse. The list is focused on the projects completed, from the funding period 2007-2013. In order to accelerate the process, the maximum level of detail required relates to Tier 2.

Romanian IM: The IM send the first project analysed and the remaining projects will be submitted in the following weeks. They explained that the delay is related to technical problems.

Italian IM: Data from the Italian IM are being provided as soon as converted into the format requested to comply to the database used in the study. Additional data are expected to be provided in the next future.

Progress Meeting with DG REGIO representatives

A progress meeting took place in Bruxelles on the 19th October. The aim of the meeting was to present the quality of the data available, the progresses on the data manipulation and analysis, to show the preliminary results and to agree with the EC on the next steps.

In addition, the process to develop the **REGIO Rail Unit Cost Database** was presented to the Commission. The database includes data extrapolate from DG REGIO and INEA databases, received from the stakeholders and retrieved from the PwC literature review.

The main steps carried out for the **data analysis** (as in the following paragraph) are:

- data normalisation and refinement process;
- data analysis;
- identification of the "base infrastructure" and calculation of unit costs.

The **preliminary results** on the estimated costs for the "base infrastructure" were introduced to the commission both for conventional and high speed lines. A preliminary analysis on the discrepancies between the estimated and final costs was presented as well, highlighting the possible causes for the cost overrun/savings.

Furthemore, the preliminary results were benchmarked with the results of the other relevant studies in the literature, showing a considerable improvement both on the data sample considered and the

cost ranges obtained.

The preliminary results of the analyses made running a regression analysis and attribute selection methodology were presented. The number and quality of the data, nonetheless represented a significant **limitation** to obtain useful and reliable results.

Eventually, the **following steps** to undertake were discussed and agreed with the Commission:

- The data normalisation approach will be refined;
- The base infrastructure cost ranges will be investigated;
- The impact of the major civil engineering structure on the construction costs will be exanimated.

Data manipulation and analysis

The **categorisation process** continued on the projects in the unique REGIO Rail Unit Cost Database. The categorisation of the projects was refined throughout the analysis, with the identification of six macro categories: base, rehabilitation and upgrading, signalling and electrification, civil structures, stations and others.

The second step consisted on **data normalisation** on produced price index and to elaborate the preliminary statistics for each category at European level and, where possible, at national level. This held the basis for more detailed analysis and to identify the main outliers.

The **out-of-range observations** were investigated in detail whether to:
- Identify the reason for inconsistencies;
- Looking for additional details on project costs and technical characteristics;
- Asking the stakeholders for clarifications and/or additional data.

The consistency of the data within each project category is fundamental to compare them and to reduce the unit cost ranges.

The first result of the analysis is the configuration of a base cost. The base infrastructure cost was defined as the cost to construct a new high speed or conventional line. In terms of the cost breakdown structure, it corresponds to the cost of the earthworks, the minor/hydraulic structures, the interfaces and interferences and the equipment.

The result of the data analysis laid the basis for the identification of the main costs drivers and for more detailed analysis.

Eventually, a multivariate linear regression analysis has been performed through OLS and clustered regression methods. The number of observation and the quality of data and available are not sufficient to guarantee high correlations and therefore to derive a general rule which links the technical characteristics and the cost of railway infrastructures.

Case Study on Tunnels

A meeting with professior Coli, from the University of Florence was held on 20th Obctober to define the structure of the case study, the main features and element to address, the technical and technological elements defining tunnels. A set of cases are identified to define the boundaries of the case study and populate it with a wide range of tunnel types as well as factors that impact on tunnel construction costs. The detailed data on tunnels are being gathered and at the same time the cases are being analysed.

Second Progress Report

The second progress report is currently under completion and it has been agreed to send it to the Commission within the 7th of November, including the progress of the analyses also in light of the meeting held on 19th October. It is based on the structure sent and agreed with the Commission, that will also be used in the final report.

-

2.	Next steps
	- In the coming weeks, a comprehensive data analysis will be performed in order to determine refined cost ranges per Tier, per category and where possible at Country level proceeding through bottom-up statistical approaches (frequency based on smaller clusters) - Although the main focus will be given to the bottom up statistical approaches, the multivariate analysis will be further investigated on the data refined and clustered; - The stakeholders consultation will continue in the following weeks, with follow ups and emails in order to complete collection of the quantitative data and to clarify the reasons for inconsistencies within the data already provided; - The case study on tunnels is being performed and additional information are being collected on specific cases, to present the highest validity and usability of the analysis. - The Second Progress Report will be submitted to the EC by the 7th November, as agreed in the Progress Meeting.

Addressees	**Kyriacos Ktenas (DG REGIO – F.1)**
Author	**PwC**
Project Name	**Assessment of unit costs (standard prices) of rail projects (Capital Expenses)**
Subject	**PROGRESS NOTE 17 November 2017**

3.	State of play and key progress from the last Progress Note (October 27ᵗʰ) to November 17ᵗʰ

Second Progress Report

The second progress report was successfully completed and submitted to the Commission within the deadline (7ᵗʰ November 2017). The report was developed with the same structure of the Final Report, as agreed with the Commission. The information included in the report will be updated as the analysis on the data available will be concluded.

Progress Meeting with the SGUCR Members

A progress meeting took place in Bruxelles on the 17ᵗʰ November. The aim of the meeting was to present the data collected in form of the REGIO Rail Unit Cost Database and the preliminary results that were obtained from the elaborations.

The different approaches undertaken were presented in order to:

- Highlight the main limitations of the "Big data approach";

- Discuss the benefits and refine the Bottom-up approach.

At the end, the following steps were presented and agreed with the Commission.

Data manipulation and analysis

The data analysis was focused on Tier 2 and Tier 3 level. In particular, **descriptive statistics** were elaborated on smaller portions of the RRUC Database to determine unit cost ranges for each rail infrastructure component.

The refinement of the **categorisation process** and the investigation of the **out-of range observations** continued throughout the analysis, in order to ensure the consistency of the data, and the refinement of the unit cost ranges for each category.

Once a satisfying cost range was identified for the single cost components, the total cost of the infrastructure was analysed by summing up the costs of the single components, through the so called, **bottom-up approach**. The approach allowed to obtain significant unit cost results with the data available.

Although the data analysis was mainly focused on the bottom-up approach, the **multivariate linear regression** was further investigated. The data available was clustered in different subcategories, based on the relevant characteristics affecting the unit cost. The regression analysis was performed on the smaller clusters identified through the OLS method, in order to increase the correlation within the variables and to reduce the heteroskedasticity, with respect to the previous analyses performed. The multivariate linear regression analysis still did not produce reliable results.

Case Study on Tunnels

The case study on the tunnels was developed. The involvement of the experts and the strong literature review performed, allowed for identifying the technical parameters and environmental factors which have a major influence on the construction costs, including: length, cross-sectional area, regulatory requirements, location, material cost and labour cost.

The technical and cost data on the tunnels included in the analysis were gathered and a preliminary analysis was performed to integrate the qualitative information with quantitative results. At the end, the unit cost range for each typology of tunnel will be determined through the analysis.

Country Reports

The Country reports have been compiled with qualitative information provided by the qualitative stakeholder consultation and integrated with desk research on additional external cost factors.