The Road Maintenance Planning Game

Game design and first results

Scharpff, Joris; Schraven, Daan; Volker, Leentje; Spaan, Matthijs; de Weerdt, Mathijs

Publication date
2019

Document Version
Other version

Citation (APA)
Scharpff, J., Schraven, D., Volker, L., Spaan, M., & de Weerdt, M. (2019). The Road Maintenance Planning Game: Game design and first results.

The Road Maintenance Planning Game

Game design and first results

Joris Scharpff[a,*], Daan Schraven[b], Leentje Volker[c],
Matthijs T.J. Spaan[a] and Mathijs M. de Weerdt[a]

[a]Delft University of Technology, Faculty of Electrical Engineering, Mathematics and Computer Science
[b]Delft University of Technology, Faculty of Civil Engineering and Geosciences
[c]University of Twente, Faculty of Engineering Technology

Abstract

This white paper describes the Road Maintenance Planning game, a game that simulates planning, coordination and execution of maintenance projects in the domain of infrastructural maintenance. In particular, the game models the dynamic contracting procedure of Volker et al. (2014), an innovative way of contracting public works to a team group of service providers. Foremost, this paper describes the game design, its practical set-up and the methodology for collecting data from gaming sessions so that future researchers can make use of the game. Additionally, this white paper includes a complete overview of the first empirical results obtained from 7 gaming sessions as part of the research of Scharpff et al. (2019). The source code and design documents can be found on GitLab[1] and may be used for academic purposes only.

This paper is organised as follows. Chapter 1 describes the game design and setup in detail, including a level playing field game model. Chapter 2 describes the methodology to gather both qualitative as well as quantitative data regarding the decision making, performance and context of players. Chapter 3 presents the full results of 7 gaming sessions that were performed in the context of the study by Scharpff et al. (2019). Finally, Chapter 4 validates the correctness of the game design by correlating the played strategies to the outcomes, using the results of the aforementioned gaming sessions.

*Email: j.c.d.scharpff@tudelft.nl
[1]`https://gitlab.com/jscharpff/maintenance-planning-game`

1 | Game Design

The Road Maintenance Planning game is designed as a level playing field so that every player faces the same set of challenges. The network that has been used in the sessions is depicted by the fully-connected pentagram of Figure 1.1, modelling a network of six fictive cities (orange) connected by a total of 45 road segments. Every player is given a maintenance portfolio consisting of four projects, corresponding to the road segments assigned through tender, that need to be serviced within the period of 1 year (assumed equal to exactly 52 weeks). The assignment of road segments to players is indicated through the colours of the segments, e.g. the red player is responsible for the maintenance of segments B12-A, B12-B, B12-C and B1-C. Furthermore, for each of these segments, the players are given four alternative ways to perform the maintenance, `LOW COST`, `LOW TTL`, `NO RISK` and `FAST`, that impact the objectives of a player in different ways. Whereas for instance the `LOW COST` alternative has low execution costs but a higher ttl, the `NO RISK` method ensures that the maintenance will not delay and thus ensuring a robust execution with predictable costs and ttl. How and when to schedule their maintenance within the 52 weeks is entirely up to the player's preference and strategy. Note that a player may also decided not to perform maintenance of a segment, at a penalty of $1/3$ times its agreed price upon completion. The design of these alternatives is listed in Table 1.1 for the pink player.

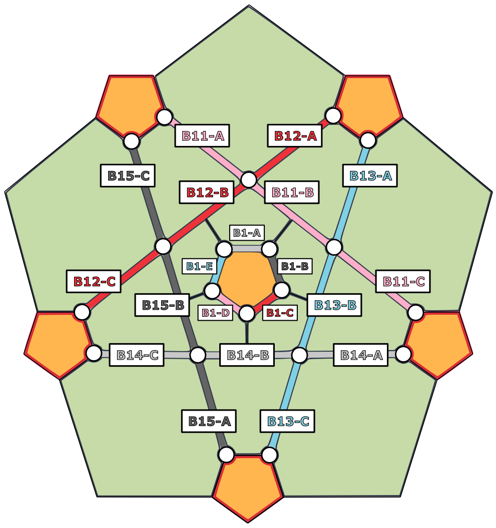


Figure 1.1: *Illustration of the level playing field network that is used in the gaming sessions.*

Segment	Cost (€)	TTL (h/w)	Duration (weeks)	Delay (weeks)	Delay probability
B1-D					
`LOW TTL`	3,184,500	8,000	10	2	33 %
`LOW COST`	2,026,500	12,000	9	4	33 %
`NO RISK`	3,184,500	10,000	9	-	-
`FAST`	2,895,000	12,000	6	2	33 %
B11-A					
`LOW TTL`	5,778,000	1,760	9	2	33 %
`LOW COST`	3,676,500	2,640	8	4	33 %
`NO RISK`	5,778,300	2,200	8	-	-
`FAST`	5,253,000	2,640	6	2	33 %
B11-B					
`LOW TTL`	4,618,500	3,040	14	2	33 %
`LOW COST`	2,938,500	4,560	13	4	33 %
`NO RISK`	4,618,350	3,800	13	-	-
`FAST`	4,198,500	4,560	9	2	33 %
B11-C					
`LOW TTL`	1,848,000	1,120	6	1	33 %
`LOW COST`	1,176,000	1,680	5	2	33 %
`NO RISK`	1,848,000	1,400	5	-	-
`FAST`	1,680,000	1,680	4	1	33 %

Table 1.1: *Maintenance portfolio of the pink player listing the available alternatives and their impact on the joint plan in terms of cost (euros), ttl (hours per week) and (potential) duration in weeks.*

The alternatives listed in Table 1.1 are for the pink player but, due to the level playing field design, the values in this table apply also to the other players albeit for different road segments. Furthermore, the total cost and ttl of a method is presented to the player as an expected value estimate based upon the delay probability, the extension of duration and regular cost. That is, the expected cost of an alternative is computed informally by the formula:

$$\mathbb{E}_{cost} = (1 - delay\ probability) \times cost + (delay\ probability) \times cost \times \frac{delay\ duration}{regular\ duration} \tag{1.1}$$

e.g. the `LOW COST` alternative for segment B1-D has an expected cost of $(1-0.33)\times 3,184,500 + 0.33 \times 3,184,500 \times \frac{10+2}{10} = 3,394,677$. For the ttl costs a similar estimate is applied, but ttl is also affected by the network due to the presence of the other players. This is described in the next paragraph. Note that the *actual* cost of a method follows from the execution of the game, when outcomes are realised at random for every method. In other words, either a task delays or does not due as a result of a stochastic process, and the cost to the player is either the listed cost or the cost plus delay time (e.g. $cost \times \frac{delay\ duration}{regular\ duration}$).

Network Model The traffic model underlying the network of the game is a factor-based model that represents the effect of maintenance on the road network throughput as sums of linear functions over averaged 'normal' traffic conditions. For every method m^r for road segment r and every other road segment $r' \neq r$ there exists a function of the form $f(m^r, r') \times \ell(r,t)$ that gives the increased traffic at road r due to the execution of m^r_x at time t concurrent with any maintenance on road r'. Notice that normally, in the absence of maintenance, the traffic time lost would have been $\ell(r,t)$. As only the increase in ttl with respect to the idle situation is interesting to measure, factors $f \in \mathbb{R}^+$ are used. Each player is charged an equal share of the total ttl increase caused by their joint plan, or:

$$\ell_i(\mathbf{y}, t) = \frac{\sum_{m^{r'} \neq m^r_i \in \mathbf{y}(t)} f(m^r_i, r') \times \ell(r,t)}{|\mathbf{y}(t)|}$$

In this formula, $\mathbf{y}(t)$ is shorthand notation for all methods at time t in joint plan $\mathbf{y}$ and $m^r_i \in \mathbf{y}_i(t)$ the alternative chosen by player i for road segment r at time t. Note that $\mathbf{y}(t)$ contains at most one method of each player for every time t as they cannot perform two projects at once. Furthermore, the model is anonymous as the ttl costs are symmetric with respect to the players, i.e. $\ell_i(\mathbf{y}, t) = \ell_j(\mathbf{y}, t)$ for every $i, j \in \mathbf{N}$.

The idle traffic time lost values $\ell(r,t)$ are given in Table 1.2, where they are grouped in 4 week intervals. The ttl increase factors $f(m_x, r)$ are given in Table 1.3. Once more only the network matrix for the pink portfolio is shown in the latter table; the matrix for the other players is exactly the same but for the segment names.

	Week numbers												
Segments	**1-4**	**5-8**	**9-12**	**13-16**	**17-20**	**21-24**	**25-28**	**29-32**	**33-36**	**37-40**	**41-44**	**45-48**	**49-52**
`B1-A, B1-B, B1-C, B1-D, B1-E`	9,640	11,148	14,649	15,134	15,134	8,886	15,080	8,348	8,132	15,565	14,326	19,173	18850
`B2-A, B3-A, B4-A, B5-A, B6-A`	1,687	1,951	2,564	2,648	2,648	1,555	2,639	1,461	1,423	2,724	2,507	3,355	3299
`B2-B, B3-B, B4-B, B5-B, B6-B`	16,630	19,231	25,270	26,106	26,106	15,329	26,013	14,400	14,029	26,849	24,713	33,074	32516
`B11-A, B12-A, B13-A, B14-A, B15-A`	1,938	2,241	2,945	3,042	3,042	1,786	3,031	1,678	1,635	3,129	2,880	3,854	3789
`B11-B, B12-B, B13-B, B14-B, B15-B`	3,516	4,066	5,343	5,519	5,519	3,241	5,500	3,045	2,966	5,677	5,225	6,993	6875
`B11-C, B12-C, B13-C, B14-C, B15-C`	1,938	2,241	2,945	3,042	3,042	1,786	3,031	1,678	1,635	3,129	2,880	3,854	3789
`B21-A, B22-A, B23-A, B24-A, B25-A`	658	761	999	1,032	1,032	606	1,029	569	555	1,062	977	1,308	1286
`B21-B, B22-B, B23-B, B24-B, B25-B`	658	761	999	1,032	1,032	606	1,029	569	555	1,062	977	1,308	1286
`M1-A, M2-A, M3-A, M4-A, M5-A`	4,243	4,906	6,447	6,660	6,660	3,911	6,637	3,674	3,579	6,850	6,305	8,438	8296

Table 1.2: *TTL per week for each row segment, grouped per set of equivalent segments.*

	B1-A	B1-B	B1-C	B1-D	B1-E	B2-A	B2-B	B3-A	B3-B	B4-A	B4-B	B5-A	B5-B	B6-A	B6-B
B1-D															
`LOW TTL`	.302	.302	.302	.603	-	.603	-	-	-	-	-	-	-	-	-
`LOW COST`	.452	.452	.452	.905	.905	.905	-	-	-	-	-	-	-	-	-
`NO RISK`	.377	.377	.377	.754	.377	.754	-	-	-	-	-	-	-	-	-
`FAST`	.452	.452	.452	.905	.905	.905	-	-	-	-	-	-	-	-	-
B11-A															
`LOW TTL`	-	-	-	-	-	-	.330	.660	-	.330	-	-	-	-	-
`LOW COST`	-	-	-	-	-	-	.495	.990	-	.495	-	-	-	-	-
`NO RISK`	-	-	-	-	-	-	.412	.825	-	.412	-	-	-	-	-
`FAST`	-	-	-	-	-	-	.495	.990	-	.495	-	-	-	-	-
B11-B															
`LOW TTL`	.314	.314	.314	-	-	-	-	.314	-	.314	-	.314	-	-	-
`LOW COST`	.471	.471	.471	-	-	-	-	.471	-	.471	-	.471	-	-	-
`NO RISK`	.393	.393	.393	-	-	-	-	.393	-	.393	-	.393	-	-	-
`FAST`	.471	.471	.471	-	-	-	-	.471	-	.471	-	.471	-	-	-
B11-C															
`LOW TTL`	-	-	-	-	-	-	-	-	-	.210	.210	.420	-	-	-
`LOW COST`	-	-	-	-	-	-	-	-	-	.315	.315	.630	-	-	-
`NO RISK`	-	-	-	-	-	-	-	-	-	.262	.262	.525	-	-	-
`FAST`	-	-	-	-	-	-	-	-	-	.315	.315	.630	-	-	-

	B11-A	B11-B	B11-C	B12-A	B12-B	B12-C	B13-A	B13-B	B13-C	B14-A	B14-B	B14-C	B15-A	B15-B	B15-C
B1-D															
`LOW TTL`	-	-	-	-	-	-	-	-	-	-	-	.302	-	-	.302
`LOW COST`	-	-	-	-	-	-	-	-	-	-	-	.452	-	-	.452
`NO RISK`	-	-	-	-	-	-	-	-	-	-	-	.377	-	-	.377
`FAST`	-	-	-	-	-	-	-	-	-	-	-	.452	-	-	.452
B11-A															
`LOW TTL`	-	-	.660	.330	.660	.660	.330	.330	.330	-	-	-	-	-	.330
`LOW COST`	-	.990	.990	.495	.990	.990	.495	.495	.495	-	-	-	-	-	.495
`NO RISK`	-	.412	.825	.412	.825	.825	.412	.412	.412	-	-	-	-	-	.412
`FAST`	-	.990	.990	.495	.990	.990	.495	.495	.495	-	-	-	-	-	.495
B11-B															
`LOW TTL`	-	.628	-	.628	.628	.628	.314	.628	.628	.314	.314	.314	-	-	.314
`LOW COST`	-	.942	.942	.942	.942	.942	.471	.942	.942	.471	.471	.471	-	-	.471
`NO RISK`	-	.785	.393	.785	.785	.785	.393	.785	.785	.393	.393	.393	-	-	.393
`FAST`	-	.942	.942	.942	.942	.942	.471	.942	.942	.471	.471	.471	-	-	.471
B11-C															
`LOW TTL`	-	.210	.420	-	.210	.210	-	.420	.420	.210	.210	.210	-	-	-
`LOW COST`	-	.315	.630	.630	.315	.315	-	.630	.630	.315	.315	.315	-	-	-
`NO RISK`	-	.262	.525	.262	.262	.262	-	.525	.525	.262	.262	.262	-	-	-
`FAST`	-	.315	.630	.630	.315	.315	-	.630	.630	.315	.315	.315	-	-	-

	B21-A	B21-B	B22-A	B22-B	B23-A	B23-B	B24-A	B24-B	B25-A	B25-B	M1-A	M2-A	M3-A	M4-A	M5-A
B1-D															
`LOW TTL`	-	-	-	-	-	-	-	-	-	-	-	.302	-	.302	.603
`LOW COST`	-	-	-	-	-	-	-	-	-	-	-	.452	-	.452	.905
`NO RISK`	-	-	-	-	-	-	-	-	-	-	-	.377	-	.377	.754
`FAST`	-	-	-	-	-	-	-	-	-	-	-	.452	-	.452	.905
B11-A															
`LOW TTL`	.330	.330	-	-	-	-	-	-	-	-	.330	.330	.330	-	-
`LOW COST`	.495	.495	-	-	-	-	-	-	-	-	.495	.495	.495	-	-
`NO RISK`	.412	.412	-	-	-	-	-	-	-	-	.412	.412	.412	-	-
`FAST`	.495	.495	-	-	-	-	-	-	-	-	.495	.495	.495	-	-
B11-B															
`LOW TTL`	.314	-	-	-	-	-	-	-	-	-	-	.314	.628	.314	-
`LOW COST`	.471	-	-	-	-	-	-	-	-	-	-	.471	.942	.471	-
`NO RISK`	.393	-	-	-	-	-	-	-	-	-	-	.393	.785	.393	-
`FAST`	.471	-	-	-	-	-	-	-	-	-	-	.471	.942	.471	-
B11-C															
`LOW TTL`	-	-	-	-	.210	.210	-	-	-	-	-	-	.210	.210	-
`LOW COST`	-	-	-	-	.315	.315	-	-	-	-	-	-	.315	.315	-
`NO RISK`	-	-	-	-	.262	.262	-	-	-	-	-	-	.262	.262	-
`FAST`	-	-	-	-	.315	.315	-	-	-	-	-	-	.315	.315	-

Table 1.3: *TTL factor matrix for every combination of method and road segment. Every cell contains the first 3 decimals of factor* $f(m^r, r')$*, i.e.* $f(\texttt{LOW COST}^{B1\text{-}D}, B1\text{-}A) = 0.302$*. A dash indicates a factor zero.*

Gameplay The game is played in two phases: a planning phase and an execution phase. The planning phase is by far the most important phase. In this phase players develop and coordinate their maintenance plans, while receiving information about the plans of their competitors. The execution phase merely simulation the execution of the maintenance plans that have been developed as a result of the planning phase and is only meant to determine actual revenues for winner determination and “player excitement”. The latter phase is explained later when discussing the winning criteria.

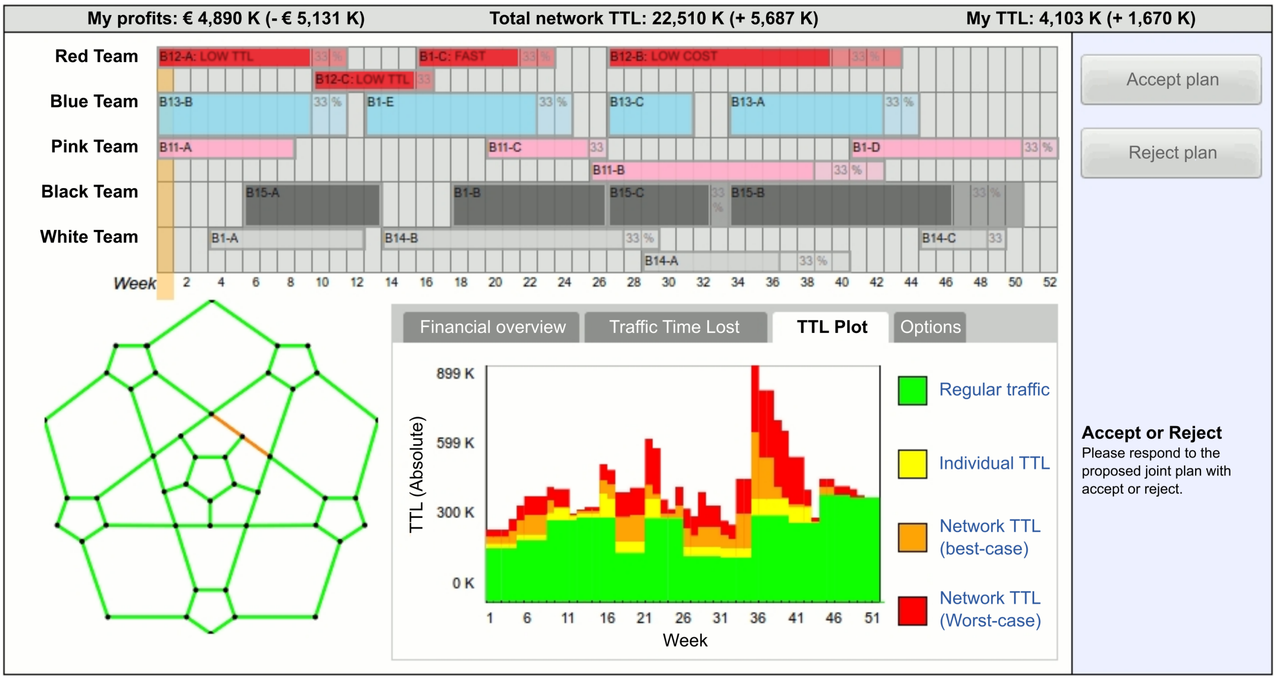


Figure 1.2: *The user interface for the players: a Gantt chart that shows the current maintenance plans, a network that shows traffic interactions and insights into the impact of the current plan.*

In the planning phase, each of the players are burdened with the task of developing a plan for the maintenance tasks in their assigned portfolio. On a Tablet PC that matches the colour of their player they are shown an interface similar to Figure 1.2 to plan their maintenance. For every project in their portfolio, a player has to decide how and when they will do the work by choosing one of the four methods described earlier and positioning it on the Gantt-like diagram in the top of the screen. The impact of their decisions in terms of cost and ttl is then computed and displayed in the bottom right table. Network relations are shown on the left to illustrate (possible) interactions with other player’s tasks.

The planning process follows a round-based procedure. In each round, every player submits its maintenance plan for its own portfolio. When all plans have been submitted, the game will combine them all into a new *joint* maintenance plan and present this plan back to the players. Players then get the option to accept or reject this new plan. If all players accept, the planning phase is over and the execution phase will start. If at least one player declines, a new planning round starts in which now all players possess information about their competitors. All last submitted plans are preserved in the interface and shown to the players. As a consequence, players can now respond to the previously made decisions of other agents. Changes to the plan, however, are not shown real-time, hence players will have to request/await a new submission of all plans to know what changes the other players have made during the new round. Observe hence that the first round is always an individual planning round: in the first planning round no plans were previously submitted and hence no information about other players’ decisions is available.

Winning the game The game can be won in two ways, mimicking to the misalignment between the contractor’s individual goal (maximum profit) and the asset manager’s global goal (maintenance with minimal traffic impact). In every session, the player that has the highest profit at the end of the game is declared the winner of that session and is rewarded a price. However the team, i.e. all players of a single session combined, that achieves the lowest expected ttl over all sessions is declared the network winner, and all members of this winning team are given a price. To incite a preference for the global goal and increase tension between both goals, the network price should be ‘more valuable’ to players than the profit price. In the gaming sessions of Scharpff et al. (2019), the profit price is a €2.50 scratch ticket whereas the ‘bigger’ team price is in the form of a €10 voucher for all network members. This will make it more interesting for players to pursue the global goal (also because of the competition against other teams) but still try to come out on top in a single gaming session, which is typically conflicting as the results of Chapter 3 show.

Winner determination is done in two ways. For a ‘fair’ comparison of all sessions, the ttl goal is measured as the expected hindrance of the final joint plan submitted by all players. In other words, this is established before going into the execution phase (but getting the lowest ttl still proved a difficult task). To determine the ‘session winner’, i.e. the contractor with the highest profit, the execution phase is started. Essentially, the only purpose of the execution phase is to simulate the projects and in particular randomly realise task delays so that the 'actual' costs and ttl become known. Although an option exists that automates this process by randomly realising delays and presenting the results back to the players, a more fun method was used during the sessions to excite players and end the session on a happy note. During the sessions, the interface was used to progress in time one week at a time. If in that week a method starts that may possibly delay, the corresponding player was asked to throw a dice to determine the outcome of that task and that outcome was fed back into the game. Indeed, this resulted in some heated endings with players in a neck-to-neck race for the price.

Practical set-up The set-up consists of one PC/Laptop that hosts the game server and five client Tablet PCs that connect to the game server as players. The game itself is developed in Java and uses the Google Web Toolkit framework for the client/server architecture and the web-based interfaces of the game. Technically there are three: the player interface as described above, the game server interface to manage the game and its progress, and a "scoreboard" interface that shows the scores and ranking of the players. The source code and network model of the Maintenance Planning Game used in this article can be found at `https://gitlab.com/jscharpff/maintenance-planning-game`.

Figure 1.3: *Four photos of the game in progress that show the setup and the interaction between players. In (a) the game is at an early stage and players are developing their individual plans on their Tablet PCs. In (b) and (c) the network planning is in full progress, where both pictures show that players in this session use a fully-coordinated decision making process to coordinate their dependencies (corresponding to the 'Very High' coordination level of Table 2.3). In particular, the whiteboard of (d) was used to regulate decision of all players. All participants in the photos agreed with the publication thereof in academic articles. The faces have been blurred to respect the privacy of the participants.*

2 | Data Gathering Methodology

This chapter describes the data gathering methodology used to validate the hypotheses of Scharpff et al. (2019) in full detail. This section is separated into three parts corresponding to the different type of measurements obtained: the agent decision preference and rationality established a priori through the questionnaire, the qualitative performance data measured from the impact of decision made in-game and the quantitative session data observed over the course of the game. Note that this section only discusses the methodology itself; the measurements obtainted from the initial gaming sessions can be found in Chapter 3

2.1 Agent decision preference and rationality

To establish the a priori decision-rationality of participants a questionnaire is used. This questionnaire poses 7 increasingly more complex decision-making scenarios from the maintenance planning domain, asking participants to rank alternatives according to their preference. Whereas the first question is relatively easy and has a 'correct' answer, i.e. the alternatives can be clearly ordered according to their ttl impact, the subsequent questions become increasingly more complex. This is due to the introduction of new factors into the decision making process such as profits, delays and the presence of other service providers. Furthermore, the alternatives are designed in such a way that no one answer is optimal in all objectives. Therefore the ranking of alternatives mostly depends on personal preference, that is, the decision rationality of the participants. The questionnaire is included below:

Dynamic Network Planning Questionnaire

Name:
Date:
Occupation and position:

Question 1

You are a service provider responsible for the maintenance of a road segment in a regional network. To this end, you have studied the impact on traffic of four possible alternatives. This results in the following congestion figures, expressed in hours of traffic time lost (TTL), caused by each alternative.

	Alternative A	Alternative B	Alternative C	Alternative D
TTL	352.000	578.000	440.000	370.000

a) Can you specify the order in which you would choose from the various alternatives? Please rank them from 1 (best) to 4 (worst).

Answer:	Rank:
Alternative A	
Alternative B	
Alternative C	
Alternative D	

b) Please motivate your ranking.

Question 2

In addition you perform cost computation, resulting in the following figures:

	Alternative A	Alternative B	Alternative C	Alternative D
TTL	352.000	578.000	440.000	370.000
Profit	€ 1.450.000	€ 2.108.000	€ 1.500.000	€ 1.739.000

a) Can you specify the order in which you would choose from the various alternatives? Please rank them from 1 (best) to 4 (worst).

Answer:	Rank:
Alternative A	
Alternative B	
Alternative C	
Alternative D	

b) Please motivate your ranking.

Question 3

The road authority decides to implement a traffic penalty payment that charges the service provider 1 euro for each hour of TTL. After some recalculation you find out that this has the following impact on your project:

	Alternative A	Alternative B	Alternative C	Alternative D
TTL	352.000	578.000	440.000	370.000
Profit	€ 1.098.000	€ 1.529.000	€ 1.060.000	€ 1.369.000

a) Can you specify the order in which you would choose from the various alternatives? Please rank them from 1 (best) to 4 (worst).

Answer:	Rank:
Alternative A	
Alternative B	
Alternative C	
Alternative D	

b) Please motivate your ranking.

Question 4

You are aware of the possibility that your project execution might be delayed and you are wondering how much that will affect the figures from before. Therefore you decide to also consider this delay in your computations:

Situation	Risk of delay		Alt. A	Alt. B	Alt. C	Alt. D
Project as planned	67%	*TTL*	352.000	578.000	440.000	370.000
		Profit	€ 1.098.000	€ 1.529.000	€ 1.060.000	€ 1.369.000
Project is delayed	33%	*TTL*	443.000	885.000	440.000	503.000
		Profit	€ 1.006.000	€ 1.223.000	€ 1.060.000	€ 1.236.000

a) Can you specify the order in which you would choose from the various alternatives? Please rank them from 1 (best) to 4 (worst).

Answer:	Rank:
Alternative A	
Alternative B	
Alternative C	
Alternative D	

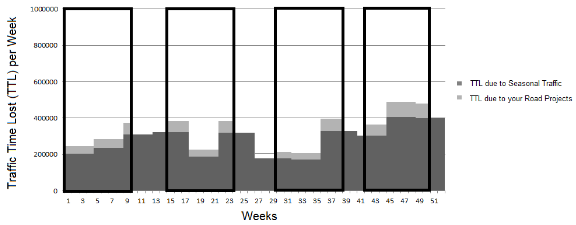
b) Please motivate your ranking.

Question 5

You have chosen four potential periods in which you can perform your project. Using quarterly figures, you determine the following prospects regarding four possible maintenance periods:

Situation	Risk of delay		Period 1	Period 2	Period 3	Period 4
Project as planned	67%	*TTL*	370.000	416.000	333.000	615.000
		Profit	€ 1.369.000	€ 1.323.000	€ 1.406.000	€ 1.124.000
Project is delayed	33%	*TTL*	503.000	571.000	493.000	809.000
		Profit	€ 1.236.000	€ 1.168.000	€ 1.246.000	€ 930.000

In addition, you also possess information regarding the TTL figures of the previous year.



a) Can you specify the order of periods in which you prefer to perform the maintenance? Please rank them from 1 (best) to 4 (worst).

Answer:	Rank:
Period 1	
Period 2	
Period 3	
Period 4	

b) Please motivate your ranking.

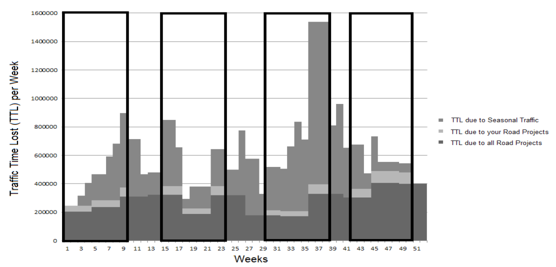
Question 6

Other service providers are also working in this region and they, in combination with your project, cause additional traffic hindrance:

Situation	Risk of delay		Period 1	Period 2	Period 3	Period 4
Project as planned	67%	*TTL*	370.000	416.000	333.000	615.000
		Profit	€ 1.369.000	€ 1.323.000	€ 1.406.000	€ 1.124.000
Project is delayed	33%	*TTL*	503.000	571.000	493.000	809.000
		Profit	€ 1.236.000	€ 1.168.000	€ 1.246.000	€ 930.000

Situation	Risk of delay		Period 1	Period 2	Period 3	Period 4
Project as planned	67%	*TTL Ind.*	370.000	416.000	333.000	615.000
		TTL Net.	990.000	2.033.000	3.966.000	1.302.000
		Profit	€ 1.171.000	€ 916.000	€ 613.000	€ 863.000
Project is delayed	33%	*TTL Ind.*	414.000	471.000	406.000	667.000
		TTL Net.	2.030.000	2.662.000	6.545.000	1.436.000
		Profit	€ 919.000	€ 736.000	€ 24.000	€ 785.000

In this table, the individual TTL denotes the TTL caused solely by your project, ignoring others. The network TTL captures the 'combined effect' of multiple service providers working concurrently in the same region. Next to this table, you are also given a plot of the TTL distribution over time.



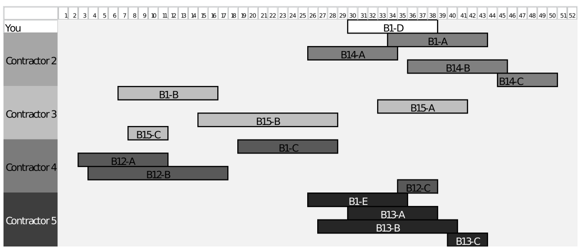
a) Can you specify the order in which you would choose from the various periods? Please rank them from 1 (best) to 4 (worst).

Answer:	Rank:
Period 1	
Period 2	
Period 3	
Period 4	

b) Please motivate your ranking.

Question 7

You have chosen the third period and you wish to reduce the incurred traffic time losses. Through communication with the other contractors, you have been able to create the following joint schedule:



a) a) If you were given the opportunity to change plans made by others, what project(s) would you modify? Can you rank the projects based on which one you would modify first (1) to last (4)?

Project:	Rank:
Contractor 2: B1-A	
Contractor 2: B14-B	
Contractor 3: B15-A	
Contractor 5: B13-A	

b) Please motivate your ranking.

Thank you very much for completing this questionnaire!

To measure the decision preference and rationality, the submitted responses are evaluated against pre-determined rankings of alternatives per question and objective. In other words, for every question n and objective m there is a ranking of objectives O_n^m that lists the alternatives in their order from best to worst. For example, the profit ordering of question 2 is the ranking $O_2^p = (4, 1, 3, 2)$ as alternative B yields the most profit and A the least. For the ttl, on the other hand, the ranking is given by $O_2^t = (1, 4, 3, 2)$ as the lowest hindrance is caused by alternative A. Note that here numerical indexes are used instead of the alphabetical index in the questionnaire, this is more convenient when computing scores. For each objective the ranking is determined according to the following rules:

- The profit ranking O^p is given by the expected profit of an alternative, ordered from highest to lowest. When the probability of delay is zero, the profit is simply the profit as listed (questions 1 to 4). In the case of potential delay, the expected profit is computed as $(1 - p) \times P_{planned} + p \times P_{delayed}$ in which p is the probability of delay (0 or 0.33) and $P_{planned}$ and $P_{delayed}$ the listed profit when respectively the maintenance is performed according to plan or a delay is encountered.

- The ttl ranking O^t is determined by the expected ttl, ordered from lowest (best) to highest (worst). Similar to profit, the ttl score is also computed by the expected ttl in the presence of delay. Moreover, in the presence of other service providers an additional ttl component 'Network' is factored into the computation, i.e. $(1 - p) \times (T_{ind,planned} + T_{net,planned}) + p \times (T_{ind,delayed} + T_{net,delayed})$ such that $T_{ind,x}$ and $T_{net,x}$ express the listed individual and network ttl for the planned and delayed scenarios. Note that it is assumed that work by other contractors does not delay (or always does) to keep the influence of the network on ttl simple.

- The risk(-aversion) ranking O^r is ordered on the expected loss of revenue due to delay from lowest to highest potential loss. In other words, the risk aversion score is the highest when the effect of delay is the least and lowest when a delay causes high revenue decreases. The formula to compute this loss is $(P_{planned} - T_{planned}) - (P_{delayed} - T_{delayed})$ due to the design of payments such that every hour of ttl incurs a cost of 1 euro (with $T_x = T_{ind,x} + T_{net,x}$).

	Profit				**TTL**				**Risk-aversion**			
Question #	**A**	**B**	**C**	**D**	**A**	**B**	**C**	**D**	**A**	**B**	**C**	**D**
1	-	-	-	-	1	4	3	2	-	-	-	-
2	4	1	3	2	1	4	3	2	-	-	-	-
3	3	1	4	2	1	4	3	2	-	-	-	-
4	2	4	3	1	1	4	3	2	2	4	1	3
5	2	3	1	4	2	3	1	4	1	2	3	4
6	1	2	4	3	1	3	4	2	3	2	4	1
7	-	-	-	-	3	4	2	1	-	-	-	-

Table 2.1: *Rank of alternatives per question and objective from best (1) to worst (4). The entries marked as '-' indicate a no-score in that objective, due to absence of the objective.*

Using these predefined rankings per objective, Table 2.1 is obtained that contains the rankings for every alternative per question and objective. Now, given a complete questionnaire response that is composed of rankings $\mathbf{x} = (x_1, x_2, \ldots, x_7)$, such that x_i ranks the alternatives for question i from best to worst, the rankings per objective can be used to compute a relative score that expresses how the participant performs in each of the objectives. This relative score is termed the *player profile score* and is an indication for the decision-making preference of the participants. To compute the profile scores, first the objective rankings of Table 2.1 are converted into weights for each alternative. Then, the submitted ranking of alternatives for each question is aggregated into a single score using the multi-criteria decision-making scoring of (Roszkowska 2013, Triantaphyllou 2013).

The weighting of alternatives is performed according to the rank-order centroid (ROC) formula proposed initially by Barron and Barrett (1996). ROC is often used in decision making theory when the relative rank ordering is known but no meaningful quantitative information is available about the alternatives. It has the property of minimising the maximum error of each weight and typically generates weight vectors that are comparable to those produced by panels of subject matter experts. Assuming that the four weights are uniformly distributed, the weight for each alternative is computed as its expected value by $\mathbb{E}(w_j) = 1/n \sum_{k=j}^{n} 1/k$, where w_j is the weight for the alternative at rank position j. In the case of four alternatives this yields the weight vector (.521, .271, .146, .063), ordered from best to worst alternative. By combining the relative rank ordering with the rank weights, the weighted rank score S_i^m per attribute m is determined for every problem i of the questionnaire by simply replacing the ranks $j \in [1, 4]$ of ordering O_i^m by their respective rank weights w_j. Thus, the ordering $O_5^p = (2, 3, 1, 4)$ of alternatives of question 5 with respect to expected profits becomes the score vector $S_5^p = (.271, .146, .521, .063)$. Similar substitutions can be performed to generate all alternative weights.

The weighted rank scores enable to measuring and comparing questionnaire responses on a quantified scale. Again a single, complete questionnaire response is denoted by $\mathbf{x} = (x_1, x_2, \ldots, x_7)$, with x_i being the ranking of alternatives for question i as submitted by the participant. As before, a ranking is a vector that for every alternative specifies the preferred order from best (1) to worst (4). Given a ranking x_i for question i, the (unscaled) profile score for objective m is then computed by $q_m(x_i) = \sum_{k=1}^{4} (5 - x_{i,k}) \times S_{i,k}^m$, such that $5 - x_{i,k}$ ensures that the first ranked alternative has a weight of

4 and the least preferred option gets a weight of 1. Given a questionnaire response $\mathbf{x}$, the preference score of a participant for objective m is the normalised sum of scores over all questions:

$$\hat{Q}_m(\mathbf{x}) = \frac{\sum_{i=1}^{|\mathbf{x}|} q_m(x_i) - Q_m^{min}}{Q_m^{max} - Q_m^{min}} \tag{2.1}$$

where Q_m^{min} and Q_m^{max} are respectively the minimum and maximum attainable scores for objective m computed over all possible rankings of alternatives. Then from the preference scores over all objectives, the actual profile score can be computed as its relative importance using the formula:

$$Q_m(\mathbf{x}) = \frac{\hat{Q}_m(\mathbf{x})}{\sum_{k\in\{p,t,r\}} \hat{Q}_k(\mathbf{x})} \tag{2.2}$$

Finally, the *rationality* of the questionnaire responses is measured in terms of their distances to the closest Pareto-optimal score and closest minimum score, also known as the normalised Pareto distance (Rousis 2011). Given a complete questionnaire score $\mathbf{q} = (Q_p(\mathbf{x}), Q_t(\mathbf{x}), Q_r(\mathbf{x}))$ with the symbols p for profit, t for ttl and r for risk-aversion, decision rationality is then expressed as the Euclidean distance of $\mathbf{q}$ to the closest Pareto-optimal score $\mathbf{b}$ inversely related to the sum of Euclidean distances to score $\mathbf{b}$ and closest lowest score $\mathbf{w}$:

$$\theta(\mathbf{q}) = 1 - \frac{||\mathbf{b} - \mathbf{q}||}{||\mathbf{b} - \mathbf{q}|| + ||\mathbf{w} - \mathbf{q}||} \tag{2.3}$$

Here, the closest Pareto scores $\mathbf{b}$ and $\mathbf{w}$ are determined by checking the distance from $\mathbf{q}$ to all other scores. The sets of the worst and best Pareto scores for the questionnaire responses are computed using a simple Java program that can be found in the code base on GitLab at `https://gitlab.com/jscharpff/maintenance-planning-game`.

2.2 Player/team Strategy

The in-game actions are scored based on their impact on each of the objectives as discussed in Chapter 1. Each maintenance alternative available to the players in the game is attributed scores for profit, ttl and risk-aversion, and they are ranked from best to worst, similar to the ranking of the previous section. These rankings are shown in Table 2.2.

Method	G_p	G_t	G_r
LOW TTL	4	1	2
LOW COST	1	4	4
NO RISK	3	3	1
FAST	2	2	3

Table 2.2: *Rank of each maintenance method per objective from best (1) to worst (2).*

From these rankings the *strategy score* for a player is computed from his/her submitted maintenance plan as follows. For a single player $i \in \mathbf{N}$, a maintenance plan is given by

$$\mathbf{y}_k = (\langle m_1^i, t_1^i\rangle, \langle m_2^i, t_2^i\rangle, \langle m_3^i, t_3^i\rangle, \langle m_4^i, t_4^i\rangle)$$

such that each $\langle m_k^i, t_k^i\rangle$ represents the chosen alternative m_k^i and start time t_k^i of maintenance task k. The played preference score G_p of player i is then

$$\hat{G}_p(\mathbf{y}_i) = \frac{\sum_{m_k^i \in \mathbf{y}_i} g_p(m_k^i) - G_p^{min}}{G_p^{max} - G_p^{min}} \tag{2.4}$$

in which $g_p(m_k^i)$ is the (non-normalised) profit score for method m_k^i of task k and G_{min}^p and G_{max}^p denote respectively the minimum and maximum profit scores attainable in game. As with the profile scores of Section 2.1, the profit strategy score G_p, ttl strategy score G_t and risk-aversion strategy score G_r are computed relative to the other preferences, thus:

$$G_m(\mathbf{y}_i) = \frac{\hat{G}_m(\mathbf{y}_i)}{\sum_{k\in\{p,r,t\}} \hat{G}_k(\mathbf{y}_i)} \tag{2.5}$$

and strategy scores of a session are aggregated using the average strategy score over all teams.

The performance of players with respect to game outcomes is also scored. To this end, the expected profit, traffic time lost and performance are measured as a function of the former two, similar to the model of Scharpff et al. (2013). The expected profit of a player i given its plan $\mathbf{y}_i$, denoted by $P_i(\mathbf{y}_i)$, is defined as the expected reward of completing work minus the expected costs thereof, or

$$P_i(\mathbf{y}_i) = \sum_{m_k^i \in \mathbf{y}_i} \Big(W(m_k^i) \; - \sum_{t=t_k^i}^{t_k^i + d(m_k^i)} C(m_k^i, t) - p(m_k^i) \sum_{t=t_k^i}^{t_k^i + \hat{d}(m_k^i)} C(m_k^i, t) \Big) \tag{2.6}$$

such that $d(m_k^i)$ and $\hat{d}(m_k^i)$ denote respectively the regular and extended maintenance period, the latter only applies when the task is delayed with probability $p(m_k^i)$. Furthermore, $W(m_k^i)$ is the fixed, contracted reward received upon completion of the task associated with m_k^i (thus independent from the chosen method) and $C(m_k^i, t)$ the maintenance cost of performing method m_k^i at time t.[1]

For the traffic time lost $T_i(\mathbf{y})$ caused by player i, given *joint* plan $\mathbf{y} = \bigcup_{i \in \mathbf{N}} \mathbf{y}_i$ with the set of players $\mathbf{N} = \{1, 2, \ldots, 5\}$, a similar expected value computation is made. Notice that for the computation of ttl a joint plan is required as concurrent maintenance can have super-linear impact on traffic. The ttl model of the game is defined through a function $\ell_i(\mathbf{y}, t)$ (Chapter 1) that returns the ttl caused by player $i \subseteq \mathbf{N}$ at time t when joint plan $\mathbf{y}$ is executed. Therefore, the total ttl caused by an individual player i is given by[2]

$$T_i(\mathbf{y}) = \sum_{m_k^i \in \mathbf{y}_i} \Big(\sum_{t=t_k^i}^{t_k^i + d(m_k^i)} \ell_i(\mathbf{y}, t) + p(m_k^i) \sum_{t=t_k^i}^{t_k^i + \hat{d}(m_k^i)} \ell_i(\mathbf{y}, t) \Big) \tag{2.7}$$

With the aforementioned formulas the expected utility of a player i is expressed as the sum of its expected revenue minus the monetary value of the expected ttl. Consequentially, expected profit for a player i given a joint plan $\mathbf{y}$, such that $\mathbf{y}_i \in \mathbf{y}$ is the plan of player i, is given by $u_i(\mathbf{y}) = P_i(\mathbf{y}_i) - T_i(\mathbf{y})$. Finally, similar to the decision rationality of profile scores, an indication of the quality of the in-game decisions can be defined over the strategy scores. The *performance ratio* $\phi(\mathbf{y})$ for a given joint plan $\mathbf{y}$ expresses the ratio between profit and ttl:

$$\phi(\mathbf{y}) = \frac{P(\mathbf{y})}{T(\mathbf{y})} \tag{2.8}$$

and observe that this value increases either when the joint profit increases, the joint ttl decreases or both. Hence a higher performance ratio indicates a better overall outcome.

2.3 Session metadata

In addition to the qualitative measurements, two of quantitative measurements is also defined that characterise the sessions in terms of coordination and familiarity between players in a single session. These characteristics have not been measured in terms of some absolute figure; in stead they are determined based on observations made prior to and during the gaming sessions. The coordination level is determined during the session itself and is set to the highest degree of collaboration observed. For instance, in a session where players use plenary sessions to coordinate decision but do not apply governed decision making, the coordination level is set to 'Medium'. Finally, the cohesion level is determined mostly a priori and confirmed with the group at the start of the session. All the qualitative measurements are defined in Table 2.3.

	Description
Coordination level	
Low	Conflict-driven coordination of interactions via bilateral or trilateral negotiations
Medium	Coordination of network via democratic, plenary negotiations
High	Centralised planning that governs network decisions
Cohesion level	
Unfamiliar	Players have (had) limited to no interaction previously
Familiar	Players see and/or work with each other on a regular basis

Table 2.3: *Definition of the qualitative categories for each of the data points obtained through observation with their abbreviations and a short description.*

[1]The reward and cost functions have no player index as the underlying model is the same.

[2]No time step t is counted more than once due to the one-task-at-a-time restriction.

3 | Results from the sessions

This chapter summarises the results of all measurements taken from this session, both a priori as well as during the session. In total, 7 gaming sessions have been performed with 95 players from various ages, institutions and backgrounds. Each of the sessions is given a letter for identification purpose and the characteristics of these sessions are listed in Table 3.1.

	Company/institute Profile	#P	#Q	Category	Coordination	Cohesion
A	University, Computer Science	9	9	Students	Low	Unfamiliar
B	ICT-focused R&D Company	10	9	Engineers	Low	Familiar
C	Utility provider, mainly power	15	3	Professionals	Low	Unfamiliar
D	Dutch national road authority	17	16	Trainees / Interns	High	Familiar
E	Dutch national road authority	8	5	Trainees / Interns	Medium	Familiar
F	AM Professionals Course	20	9	Professionals	Medium	Unfamiliar
G	AM and Health-care Consultants	16	9	Professionals	High	Familiar

Table 3.1: *Outline of game session characteristics, from left to right the columns are: session identifier, company/institute, number of participants, number of questionnaires reviewed, participants skill category, and the observed coordination and social cohesion of participants.*

3.1 Questionnaire Responses

From the 95 participants, 59 valid questionnaires were collected. All of these responses have been scored according to Equation 2.2 and are listed in Table 3.2 on the next page. The columns capture respectively the session name, the computed profit, ttl and risk-aversion profile scores (Q_p, Q_t and Q_r), and the decision rationality θ. The average profile score of the session is included in the bottom row, included only for reference and is not used. The rationality scores are found using Equation 2.2 of the appendix, where the optimal Pareto trade-offs have been computed using a Java program that can be found in the GitLab repository.

	Q_p	Q_t	Q_r	θ
Session A				
	0.455	0.336	0.209	0.795
	0.436	0.352	0.212	0.802
	0.481	0.274	0.244	0.855
	0.440	0.338	0.221	0.815
	0.322	0.373	0.305	0.932
	0.475	0.350	0.175	0.730
	0.465	0.321	0.213	0.773
	0.385	0.384	0.231	0.845
	0.253	0.464	0.284	0.385
avg	*0.412*	*0.355*	*0.233*	*0.770*
Session B				
	0.244	0.375	0.382	0.955
	0.459	0.343	0.198	0.789
	0.230	0.401	0.369	0.883
	0.244	0.460	0.296	0.691
	0.277	0.449	0.274	0.878
	0.375	0.375	0.250	0.897
	0.369	0.395	0.236	0.870
	0.360	0.380	0.259	0.904
	0.459	0.343	0.198	0.789
avg	*0.335*	*0.391*	*0.274*	*0.851*
Session C				
	0.379	0.367	0.254	0.913
	0.428	0.321	0.251	0.849
	0.329	0.358	0.314	0.962
avg	*0.379*	*0.348*	*0.273*	*0.908*
Session D				
	0.272	0.358	0.370	0.958
	0.401	0.345	0.254	0.255
	0.455	0.336	0.209	0.795
	0.459	0.343	0.198	0.789
	0.426	0.355	0.219	0.884
	0.398	0.375	0.227	0.855
	0.350	0.417	0.233	0.877
	0.326	0.378	0.297	0.916
	0.327	0.377	0.296	0.982
	0.381	0.384	0.235	0.853
	0.431	0.342	0.227	0.880
	0.335	0.447	0.218	0.710
	0.266	0.319	0.415	0.838
	0.347	0.400	0.253	0.907
	0.336	0.404	0.260	0.938
	0.349	0.264	0.387	0.974
avg	*0.366*	*0.365*	*0.269*	*0.838*

	Q_p	Q_t	Q_r	θ
Session E				
	0.386	0.341	0.273	0.948
	0.379	0.367	0.254	0.913
	0.420	0.313	0.267	0.815
	0.473	0.315	0.212	0.845
	0.463	0.314	0.223	0.891
avg	*0.424*	*0.330*	*0.246*	*0.882*
Session F				
	0.307	0.457	0.236	0.713
	0.432	0.260	0.308	0.862
	0.483	0.338	0.178	0.715
	0.459	0.343	0.198	0.789
	0.190	0.413	0.397	0.802
	0.325	0.431	0.243	0.887
	0.383	0.336	0.281	0.987
	0.337	0.425	0.237	0.731
	0.433	0.316	0.251	0.827
avg	*0.372*	*0.369*	*0.259*	*0.813*
Session G				
	0.380	0.355	0.266	0.933
	0.340	0.393	0.266	0.933
	0.467	0.336	0.197	0.790
	0.353	0.377	0.270	0.932
	0.311	0.438	0.252	0.879
	0.350	0.340	0.310	0.961
	0.254	0.378	0.368	0.963
	0.278	0.446	0.276	0.886
	0.394	0.347	0.260	0.922
avg	*0.347*	*0.379*	*0.274*	*0.911*

Table 3.2: *Complete overview of questionnaire profile scores, grouped per session. For each response the profit, ttl and risk-aversion profile scores are computed (resp. Q_p, Q_t and Q_r) and the decision rationality θ according to Equation 2.3.*

3.2 Session Outcomes

The in-game results are listed as a single table per game session. Each table contains multiple sub-tables, one for every round played in the game, and the listed figures are the values measured exactly when all players submitted their plan. Per round, the tables list for both the profit and ttl objectives the maximum value that can be obtained, the maximum impact of delay on that value and the expected value. For example, at the end of round 1 of session A, the Red player can potentially achieve a maximum profit of $€\ 5,425$. If the player is really unlucky and all of its activities are delayed, its profit decreases by $€\ 6,464$, resulting in a total loss of $€\ 1,039$. In expectation, however, its profit is $€\ 3,270$ which is of course much better than the worst-case scenario. The ttl columns are similar but for the fact that the figure in the delay column is *added* to the ttl figure in the case of delay. Note that the score listed in the last round is the score before the execution starts.

Session A	**Profit**	**Delay**	$E[P]$	**TTL**	**Delay**	$E[T]$
Round 1						
Black	4.878	-6.748	2.629	6.179	2.721	7.086
Blue	3.273	-3.446	2.124	6.390	1.107	6.759
Pink	2.907	-4.028	1.564	6.216	893	6.514
Red	5.425	-6.464	3.270	5.632	2.438	6.445
White	5.347	-5.647	3.465	5.158	1.441	5.638
Total	21.830	-26.333	13.052	29.575	8.600	32.442
Round 2						
Black	4.878	-6.748	2.629	5.512	2.735	6.424
Blue	5.646	-6.094	3.615	5.095	2.578	5.954
Pink	5.745	-5.796	3.813	5.799	2.139	6.512
Red	5.456	-6.279	3.363	550	2.177	1.276
White	6.183	-6.208	4.114	4.322	2.001	4.989
Total	27.908	-31.125	17.533	21.278	11.630	25.155
Round 3						
Black	4.989	-6.226	2.914	5.912	2.387	6.708
Blue	5.646	-6.094	3.615	6.264	1.711	6.834
Pink	6.449	-6.019	4.443	5.029	2.955	6.014
Red	7.242	-7.292	4.811	5.590	3.246	6.672
White	5.927	-5.010	4.257	4.816	1.854	5.434
Total	30.253	-30.641	20.039	27.611	12.153	31.662

Session B	**Profit**	**Delay**	$E[P]$	**TTL**	**Delay**	$E[T]$
Round 1						
Black	3.918	-5.986	1.923	5.766	1.946	6.415
Blue	3.696	-5.289	1.933	7.257	2.079	7.950
Pink	4.269	-6.521	2.095	6.824	2.506	7.659
Red	5.842	-7.297	3.410	6.287	3.234	7.365
White	88	-3.873	-1.203	6.934	984	7.262
Total	17.813	-28.966	8.158	33.068	10.749	36.651
Round 2						
Black	4.688	-5.636	2.809	3.240	1.158	3.626
Blue	4.089	-4.935	2.444	4.639	1.726	5.214
Pink	4.269	-6.521	2.095	3.753	2.389	4.549
Red	6.971	-6.352	4.854	4.968	2.832	5.912
White	1.519	-3.771	262	4.450	841	4.730
Total	21.536	-27.215	12.464	21.050	8.946	24.032
Round 3						
Black	5.095	-5.472	3.271	3.144	803	3.412
Blue	6.693	-6.808	4.424	5.673	2.148	6.389
Pink	6.491	-6.176	4.432	4.569	1.975	5.227
Red	7.762	-6.724	5.521	4.166	2.588	5.029
White	1.899	-4.043	551	5.123	1.144	5.504
Total	27.940	-29.223	18.199	22.675	8.658	25.561
Round 4						
Black	5.095	-5.472	3.271	3.086	1.009	3.422
Blue	7.565	-6.492	5.401	4.825	2.379	5.618
Pink	6.927	-5.927	4.951	4.591	1.820	5.198
Red	8.005	-6.469	5.849	4.465	2.277	5.224
White	1.960	-3.943	646	5.045	998	5.378
Total	29.552	-28.303	20.118	22.012	8.483	24.840

Session C	**Profit**	**Delay**	$E[P]$	**TTL**	**Delay**	$E[T]$
Round 1						
Black	6.797	-6.692	4.566	5.837	2.579	6.697
Blue	5.602	-6.532	3.425	5.659	2.404	6.460
Pink	5.040	-6.427	2.898	4.793	2.382	5.587
Red	3.377	-4.433	1.899	5.747	1.650	6.297
White	6.204	-6.810	3.934	6.429	2.697	7.328
Total	27.020	-30.894	16.722	28.465	11.712	32.369
Round 2						
Black	6.797	-6.692	4.566	5.744	2.271	6.501
Blue	5.859	-6.329	3.749	5.402	2.277	6.161
Pink	5.741	-6.449	3.591	4.058	2.000	4.725
Red	3.377	-4.433	1.899	5.271	1.650	5.821
White	6.978	-6.742	4.731	5.660	2.207	6.396
Total	28.752	-30.645	18.537	26.135	10.405	29.603
Round 3						
Black	5.479	-4.901	3.845	5.097	1.482	5.591
Blue	7.163	-6.599	4.963	5.212	2.086	5.907
Pink	6.524	-6.636	4.312	3.791	1.858	4.410
Red	3.566	-3.292	2.469	4.565	1.092	4.929
White	6.978	-6.742	4.731	4.882	2.245	5.630
Total	29.710	-28.170	20.320	23.547	8.763	26.468
Round 4						
Black	5.662	-5.224	3.921	4.711	1.820	5.318
Blue	6.855	-4.857	5.236	4.579	1.369	5.035
Pink	7.691	-6.654	5.473	4.055	2.084	4.750
Red	3.566	-3.292	2.469	4.697	1.044	5.045
White	8.100	-6.179	6.040	4.951	2.195	5.683
Total	31.874	-26.206	23.139	22.993	8.512	25.830

Session D	**Profit**	**Delay**	$E[P]$	**TTL**	**Delay**	$E[T]$
Round 1						
Black	2.582	-1.326	2.140	4.729	361	4.849
Blue	6.044	-5.722	4.137	4.461	1.515	4.966
Pink	3.825	-2.941	2.845	4.158	846	4.440
Red	7.898	-6.326	5.789	4.735	2.212	5.472
White	5.986	-4.675	4.428	5.489	1.462	5.976
Total	26.335	-20.990	19.338	23.572	6.396	25.704
Round 2						
Black	3.583	-1.326	3.141	2.657	204	2.725
Blue	6.241	-5.853	4.290	2.972	920	3.279
Pink	3.249	-2.248	2.500	3.078	486	3.240
Red	5.884	-5.033	4.206	2.541	964	2.862
White	4.010	-3.498	2.844	3.012	609	3.215
Total	22.967	-17.958	16.981	14.260	3.183	15.321
Round 3						
Black	3.702	-1.326	3.260	3.171	204	3.239
Blue	6.241	-5.853	4.290	2.774	1.110	3.144
Pink	3.011	-2.248	2.262	3.312	446	3.461
Red	6.152	-5.329	4.376	3.141	1.324	3.582
White	4.418	-3.760	3.165	3.193	678	3.419
Total	23.524	-18.516	17.352	15.591	3.762	16.845

Session E	**Profit**	**Delay**	$\boldsymbol{E[P]}$	**TTL**	**Delay**	$\boldsymbol{E[T]}$
Round 1						
Black	5.993	-4.215	4.588	4.223	973	4.547
Blue	3.678	-1.835	3.066	3.932	482	4.093
Pink	5.712	-4.779	4.119	4.084	1.173	4.475
Red	5.001	-3.865	3.713	4.291	970	4.614
White	4.512	-5.109	2.809	3.996	1.263	4.417
Total	24.896	-19.803	18.295	20.526	4.861	22.146
Round 2						
Black	5.993	-4.215	4.588	3.751	923	4.059
Blue	4.417	-1.726	3.842	3.218	403	3.352
Pink	5.568	-4.493	4.070	3.925	850	4.208
Red	5.802	-5.005	4.134	3.727	1.205	4.129
White	5.006	-5.026	3.331	3.378	892	3.675
Total	26.786	-20.465	19.964	17.999	4.273	19.423
Round 3						
Black	5.924	-4.187	4.528	3.420	1.025	3.762
Blue	4.561	-1.700	3.994	3.267	347	3.383
Pink	6.377	-4.678	4.818	3.307	1.054	3.658
Red	6.032	-5.034	4.354	3.722	983	4.050
White	5.006	-5.023	3.332	3.580	720	3.820
Total	27.900	-20.622	21.026	17.296	4.129	18.672
Round 4						
Black	5.790	-4.243	4.376	3.569	1.015	3.907
Blue	6.688	-3.910	5.385	3.355	698	3.588
Pink	6.377	-4.678	4.818	3.179	1.114	3.550
Red	6.413	-5.156	4.694	3.660	1.156	4.045
White	5.006	-5.026	3.331	3.675	826	3.950
Total	30.274	-23.013	22.603	17.438	4.809	19.041
Round 5						
Black	5.790	-4.243	4.376	3.651	1.015	3.989
Blue	6.655	-4.065	5.300	3.501	853	3.785
Pink	6.377	-4.678	4.818	3.230	1.176	3.622
Red	6.413	-5.156	4.694	3.722	1.156	4.107
White	5.006	-5.026	3.331	3.643	826	3.918
Total	30.241	-23.168	22.518	17.747	5.025	19.422
Round 6						
Black	5.790	-4.243	4.376	3.805	1.180	4.198
Blue	6.547	-3.955	5.229	3.706	764	3.965
Pink	7.406	-5.464	5.585	3.521	1.585	4.049
Red	6.413	-5.156	4.694	3.794	1.171	4.184
White	4.899	-5.077	3.207	3.681	851	3.965
Total	31.055	-23.895	23.090	18.507	5.551	20.357
Round 7						
Black	5.779	-4.348	4.330	4.715	1.297	5.147
Blue	6.796	-4.164	5.408	4.686	1.189	5.082
Pink	7.406	-5.464	5.585	4.143	1.849	4.759
Red	7.893	-5.949	5.910	5.261	1.859	5.881
White	4.899	-5.077	3.207	4.296	1.188	4.692
Total	32.773	-25.002	24.439	23.101	7.382	25.562

Session F	**Profit**	**Delay**	$E[P]$	**TTL**	**Delay**	$E[T]$
Round 1						
Black	4.008	-4.699	2.442	6.208	1.457	6.694
Blue	5.119	-6.227	3.043	5.386	2.020	6.059
Pink	4.752	-4.061	3.398	5.044	1.115	5.416
Red	5.480	-6.861	3.193	6.481	2.910	7.451
White	5.242	-5.183	3.514	5.712	1.974	6.370
Total	24.601	-27.031	15.591	28.831	9.476	31.990
Round 2						
Black	5.625	-4.619	4.085	3.323	1.378	3.782
Blue	5.420	-6.081	3.393	4.317	1.492	4.814
Pink	6.556	-4.204	5.155	3.766	963	4.087
Red	7.267	-6.324	5.159	4.912	1.488	5.408
White	4.752	-5.097	3.053	3.914	1.477	4.406
Total	29.620	-26.325	20.845	20.232	6.798	22.498
Round 3						
Black	5.625	-4.619	4.085	4.054	1.143	4.435
Blue	7.308	-6.260	5.221	3.814	1.395	4.279
Pink	6.556	-4.204	5.155	4.524	1.068	4.880
Red	7.097	-6.453	4.946	4.763	1.743	5.344
White	6.463	-5.936	4.484	4.170	1.836	4.782
Total	33.049	-27.472	23.892	21.325	7.185	23.720
Round 4						
Black	5.588	-5.216	3.849	4.911	2.200	5.644
Blue	7.308	-6.260	5.221	4.745	2.013	5.416
Pink	5.628	-4.385	4.166	5.230	1.446	5.712
Red	7.864	-6.807	5.595	5.550	2.207	6.286
White	6.720	-6.560	4.533	5.512	2.607	6.381
Total	33.108	-29.228	23.365	25.948	10.473	29.439

Session G	**Profit**	**Delay**	$E[P]$	**TTL**	**Delay**	$E[T]$
Round 1						
Black	4.415	-6.525	2.240	8.219	2.412	9.023
Blue	3.683	-6.793	1.419	6.822	2.586	7.684
Pink	4.716	-6.798	2.450	7.917	2.684	8.812
Red	3.972	-5.413	2.168	7.504	2.200	8.237
White	1.851	-5.718	-55	6.405	2.069	7.095
Total	18.637	-31.247	8.221	36.867	11.951	40.851
Round 2						
Black	6.342	-5.908	4.373	4.414	1.413	4.885
Blue	4.323	-3.642	3.109	3.870	851	4.154
Pink	5.947	-3.488	4.784	3.716	1.149	4.099
Red	6.455	-4.536	4.943	3.630	1.262	4.051
White	4.383	-5.228	2.640	3.522	1.112	3.893
Total	27.450	-22.802	19.849	19.152	5.787	21.081
Round 3						
Black	6.194	-3.813	4.923	3.909	581	4.103
Blue	5.456	-2.990	4.459	3.553	800	3.820
Pink	6.022	-3.793	4.758	4.102	570	4.291
Red	7.081	-4.437	5.602	3.804	863	4.092
White	6.216	-4.414	4.745	3.918	1.120	4.291
Total	30.969	-19.447	24.487	19.286	3.934	20.597
Round 4						
Black	6.309	-3.813	5.038	3.793	581	3.987
Blue	5.456	-2.990	4.459	3.649	724	3.890
Pink	6.022	-3.793	4.758	3.847	540	4.027
Red	7.081	-4.437	5.602	3.453	1.068	3.809
White	5.299	-4.012	3.962	3.313	939	3.626
Total	30.167	-19.045	23.819	18.055	3.852	19.339

Table 3.3 contains a summarised overview of the previous detailed session outcome listing per player. The columns R1 to R7 represent the rounds of the game. The profit and utility are the total session values in thousands of euros, the ttl is the session total in hours.

	R1	R2	R3	R4	R5	R6	R7	min	max	avg
Profit										
A	45.494	42.688	51.701					42.688	51.701	46.628
B	44.809	36.496	43.760	44.957				36.496	44.957	42.506
C	49.091	48.140	46.788	48.969				46.788	49.091	48.247
D	45.042	32.302	34.197					32.302	45.042	37.180
E	40.441	39.388	39.698	41.644	41.941	43.447	50.001	39.388	50.001	42.366
F	47.580	43.343	47.612	52.804				43.343	52.804	47.835
G	49.072	40.930	45.084	43.158				40.930	49.072	44.561
TTL										
A	32.442	25.155	31.662					25.155	32.442	29.753
B	36.651	24.032	25.561	24.840				24.032	36.651	27.771
C	32.369	29.603	26.468	25.830				25.830	32.369	28.568
D	25.704	15.321	16.845					15.321	25.704	19.290
E	22.146	19.423	18.672	19.041	19.422	20.357	25.562	18.672	25.562	20.661
F	31.990	22.498	23.720	29.439				22.498	31.990	26.912
G	40.851	21.081	20.597	19.339				19.339	40.851	25.467
Utility										
A	13.052	17.533	20.039					13.052	20.039	16.875
B	8.158	12.464	18.199	20.118				8.158	20.118	14.735
C	16.722	18.537	20.320	23.139				16.722	23.139	19.679
D	19.338	16.981	17.352					16.981	19.338	17.890
E	18.295	19.964	21.026	22.603	22.518	23.090	24.439	18.295	24.439	21.705
F	15.591	20.845	23.892	23.365				15.591	23.892	20.923
G	8.221	19.849	24.487	23.819				8.221	24.487	19.094

Table 3.3: *Outcomes per objective summed over all players in the session per round and summarised over all rounds.*

3.3 Strategy Scores

Session A	G_p	G_t	G_r
Round 1			
Black	0.786	0.143	0.071
Blue	0.563	0.188	0.250
Pink	0.222	0.389	0.389
Red	0.786	0.143	0.071
White	0.625	0.250	0.125
avg.	0.596	0.222	0.181
Round 2			
Black	0.786	0.143	0.071
Blue	0.563	0.188	0.250
Pink	0.786	0.143	0.071
Red	0.786	0.143	0.071
White	0.625	0.250	0.125
avg.	0.709	0.173	0.118
Round 3			
Black	0.786	0.143	0.071
Blue	0.563	0.188	0.250
Pink	0.786	0.143	0.071
Red	1.000	0.000	0.000
White	0.563	0.188	0.250
avg.	0.739	0.132	0.129

Session B	G_p	G_t	G_r
Round 1			
Black	0.500	0.333	0.167
Blue	0.714	0.071	0.214
Pink	0.500	0.313	0.188
Red	0.786	0.143	0.071
White	0.000	0.600	0.400
avg.	0.500	0.292	0.208
Round 2			
Black	0.400	0.400	0.200
Blue	0.714	0.071	0.214
Pink	0.500	0.313	0.188
Red	1.000	0.000	0.000
White	0.000	0.600	0.400
avg.	0.523	0.277	0.200
Round 3			
Black	0.400	0.400	0.200
Blue	1.000	0.000	0.000
Pink	0.786	0.143	0.071
Red	1.000	0.000	0.000
White	0.000	0.600	0.400
avg.	0.637	0.229	0.134
Round 4			
Black	0.400	0.400	0.200
Blue	1.000	0.000	0.000
Pink	0.786	0.143	0.071
Red	1.000	0.000	0.000
White	0.000	0.600	0.400
avg.	0.637	0.229	0.134

Session C	G_p	G_t	G_r
Round 1			
Black	1.000	0.000	0.000
Blue	0.625	0.250	0.125
Pink	0.389	0.389	0.222
Red	0.222	0.389	0.389
White	1.000	0.000	0.000
avg.	0.647	0.206	0.147
Round 2			
Black	1.000	0.000	0.000
Blue	0.625	0.250	0.125
Pink	0.389	0.389	0.222
Red	0.222	0.389	0.389
White	1.000	0.000	0.000
avg.	0.647	0.206	0.147
Round 3			
Black	0.714	0.071	0.214
Blue	1.000	0.000	0.000
Pink	0.625	0.250	0.125
Red	0.278	0.278	0.444
White	1.000	0.000	0.000
avg.	0.723	0.120	0.157
Round 4			
Black	0.714	0.071	0.214
Blue	0.714	0.071	0.214
Pink	0.786	0.143	0.071
Red	0.278	0.278	0.444
White	1.000	0.000	0.000
avg.	0.698	0.113	0.189

Session D	G_p	G_t	G_r
Round 1			
Black	0.250	0.250	0.500
Blue	0.625	0.250	0.125
Pink	0.333	0.333	0.333
Red	1.000	0.000	0.000
White	0.714	0.071	0.214
avg.	0.585	0.181	0.235
Round 2			
Black	0.250	0.250	0.500
Blue	0.625	0.250	0.125
Pink	0.200	0.350	0.450
Red	0.400	0.400	0.200
White	0.000	0.600	0.400
avg.	0.295	0.370	0.335
Round 3			
Black	0.250	0.250	0.500
Blue	0.625	0.250	0.125
Pink	0.200	0.350	0.450
Red	0.500	0.333	0.167
White	0.100	0.550	0.350
avg.	0.335	0.347	0.318

Session E	G_p	G_t	G_r
Round 1			
Black	0.563	0.188	0.250
Blue	0.300	0.300	0.400
Pink	0.375	0.375	0.250
Red	0.333	0.333	0.333
White	0.389	0.389	0.222
avg.	0.392	0.317	0.291
Round 2			
Black	0.563	0.188	0.250
Blue	0.300	0.300	0.400
Pink	0.375	0.375	0.250
Red	0.389	0.389	0.222
White	0.389	0.389	0.222
avg.	0.403	0.328	0.269
Round 3			
Black	0.563	0.188	0.250
Blue	0.300	0.300	0.400
Pink	0.389	0.389	0.222
Red	0.389	0.389	0.222
White	0.389	0.389	0.222
avg.	0.406	0.331	0.263
Round 4			
Black	0.563	0.188	0.250
Blue	0.444	0.278	0.278
Pink	0.389	0.389	0.222
Red	0.500	0.313	0.188
White	0.389	0.389	0.222
avg.	0.457	0.311	0.232
Round 5			
Black	0.563	0.188	0.250
Blue	0.444	0.278	0.278
Pink	0.389	0.389	0.222
Red	0.500	0.313	0.188
White	0.389	0.389	0.222
avg.	0.457	0.311	0.232
Round 6			
Black	0.563	0.188	0.250
Blue	0.444	0.278	0.278
Pink	0.389	0.389	0.222
Red	0.500	0.313	0.188
White	0.389	0.389	0.222
avg.	0.504	0.283	0.213
Round 7			
Black	0.563	0.188	0.250
Blue	0.563	0.188	0.250
Pink	0.389	0.389	0.222
Red	1.000	0.000	0.000
White	0.389	0.389	0.222
avg.	0.628	0.203	0.169

Session F	G_p	G_t	G_r
Round 1			
Black	0.563	0.188	0.250
Blue	0.625	0.250	0.125
Pink	0.438	0.250	0.313
Red	0.643	0.214	0.143
White	0.714	0.071	0.214
avg.	0.596	0.195	0.209
Round 2			
Black	0.563	0.188	0.250
Blue	0.625	0.250	0.125
Pink	0.563	0.188	0.250
Red	0.643	0.214	0.143
White	0.563	0.188	0.250
avg.	0.591	0.205	0.204
Round 3			
Black	0.563	0.188	0.250
Blue	0.625	0.250	0.125
Pink	0.563	0.188	0.250
Red	0.643	0.214	0.143
White	0.786	0.143	0.071
avg.	0.636	0.196	0.168
Round 4			
Black	0.714	0.071	0.214
Blue	0.625	0.250	0.125
Pink	0.563	0.188	0.250
Red	1.000	0.000	0.000
White	1.000	0.000	0.000
avg.	0.780	0.102	0.118

Session G	G_p	G_t	G_r
Round 1			
Black	1.000	0.000	0.000
Blue	0.625	0.250	0.125
Pink	1.000	0.000	0.000
Red	0.714	0.071	0.214
White	0.350	0.350	0.300
avg.	0.738	0.134	0.128
Round 2			
Black	0.786	0.143	0.071
Blue	0.444	0.278	0.278
Pink	0.563	0.188	0.250
Red	0.563	0.188	0.250
White	0.350	0.350	0.300
avg.	0.541	0.229	0.230
Round 3			
Black	0.563	0.188	0.250
Blue	0.444	0.278	0.278
Pink	0.563	0.188	0.250
Red	0.563	0.188	0.250
White	0.563	0.188	0.250
avg.	0.539	0.206	0.256
Round 4			
Black	0.563	0.188	0.250
Blue	0.444	0.278	0.278
Pink	0.563	0.188	0.250
Red	0.563	0.188	0.250
White	0.444	0.278	0.278
avg.	0.515	0.224	0.261

And, as with the outcomes, Table 3.4 summarises the strategy scores per session over all rounds of the game.

	R1	R2	R3	R4	R5	R6	R7	min	max	avg
Profit										
A	0.596	0.709	0.739					0.596	0.739	0.681
B	0.500	0.523	0.637	0.637				0.500	0.637	0.574
C	0.647	0.647	0.723	0.698				0.647	0.723	0.679
D	0.585	0.295	0.335					0.295	0.585	0.405
E	0.392	0.403	0.406	0.457	0.457	0.504	0.628	0.392	0.628	0.464
F	0.596	0.591	0.636	0.780				0.591	0.780	0.651
G	0.738	0.541	0.539	0.515				0.515	0.738	0.583
TTL										
A	0.222	0.173	0.132					0.132	0.222	0.176
B	0.292	0.277	0.229	0.229				0.229	0.292	0.256
C	0.206	0.206	0.120	0.113				0.113	0.206	0.161
D	0.181	0.370	0.347					0.181	0.370	0.299
E	0.317	0.328	0.331	0.311	0.311	0.283	0.203	0.203	0.331	0.298
F	0.195	0.205	0.196	0.102				0.102	0.205	0.175
G	0.134	0.229	0.206	0.224				0.134	0.229	0.198
Risk aversion										
A	0.181	0.118	0.129					0.118	0.181	0.143
B	0.208	0.200	0.134	0.134				0.134	0.208	0.169
C	0.147	0.147	0.157	0.189				0.147	0.189	0.160
D	0.235	0.335	0.318					0.235	0.335	0.296
E	0.291	0.269	0.263	0.232	0.232	0.213	0.169	0.169	0.291	0.238
F	0.209	0.204	0.168	0.118				0.118	0.209	0.175
G	0.128	0.230	0.256	0.261				0.128	0.261	0.219

Table 3.4: *Strategy profile scores objective averaged over all players in the session per round and summarised over all rounds.*

4 | Validation of the game model

For completeness, the correctness of the game model is expressed as an additional hypothesis, describing that the actions of the game have the intended effect on the outcome. This hypothesis is subsequently tested against the measurements obtained from the gaming sessions of the previous chapter to establish that indeed the actions of the game have their intended impact on the game outcomes. Put differently, higher ttl strategy scores should lead to lower ttl in the outcome of the game, and similarly for the other objectives. The correctness hypothesis is formulated as:

Hypothesis 1 (Action Consistency)**.** *The actions (maintenance alternatives) in the game are label consistent to their intended impact on the game outcomes for all of the objectives (a) profit, (b) ttl and (c) risk aversion.*

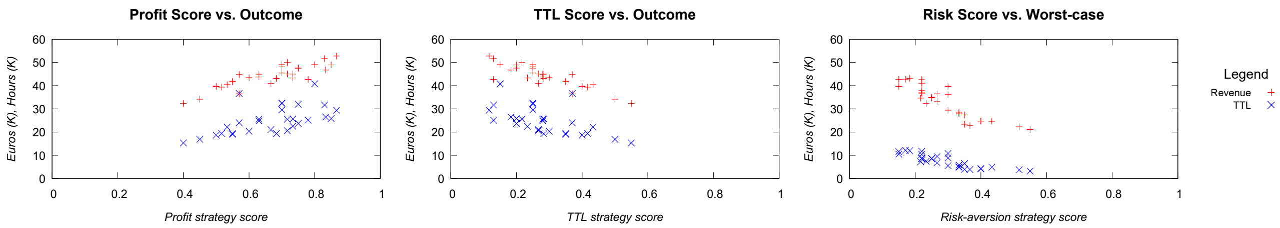


Figure 4.1: *The impact of actions on the outcome of the game: (a) and (b) show the profit and ttl strategy scores respectively versus the total expected revenue and ttl obtained from joint plans, and (c) shows the risk-aversion strategy score versus the worst case revenue loss and ttl increase.*

Figure 4.1 shows the total expected revenue and ttl of that results from each joint plan submitted at the end of a planning round versus the corresponding strategy scores (Appendix 2) in all three objectives: (a) revenue, (b) ttl and (c) risk aversion. From figures 4.1(a) and 4.1(b) a clear relation can be observed between the action choices and the effect on the outcome: a higher profit strategy score leads higher revenue and, similarly, a higher ttl strategy score is paired with a reduction in ttl. The few outliers in the 28-35K range all correspond to joint plans submitted in the initial planning round. As players do not coordinate yet in this round, they are to be expected to be much worse in terms of ttl. Coincidentally, these outliers provide a good illustration of the significance of coordination in a self-regulating network. Finally, Figure 4.1(c) shows the worst-case revenue loss and worst-case ttl increase for a joint plan as a function of the risk aversion strategy score. Here once more a strong correlation can be observed, i.e. a higher risk aversion score results in a lower worst-case revenue loss and lower worst-case ttl increase.

Acknowledgements

This work was supported by NGInfra under grant number 03.21.ALM; NWO DTC-NCAP under grant number #612.001.109; NWO VENI under grant number #639.021.336.

Bibliography

Barron, F. H. and Barrett, B. E. (1996). Decision quality using ranked attribute weights. *Management science*, 42(11):1515–1523.

Roszkowska, E. (2013). Rank ordering criteria weighting methods–a comparative overview. Technical report, Wydawnictwo Uniwersytetu w Białymstoku.

Rousis, D. (2011). *A Pareto frontier intersection-based approach for efficient multiobjective optimization of competing concept alternatives*. PhD thesis, Georgia Institute of Technology.

Scharpff, J., Schraven, D., Volker, L., Spaan, M. T., and De Weerdt, M. M. (2019). Can multiple contractors self-regulate their joint service delivery? A serious gaming experiment on road maintenance planning. *Manuscript submitted for publication*.

Scharpff, J., Spaan, M. T. J., de Weerdt, M. M., and Volker, L. (2013). Planning under uncertainty for coordinating infrastructural maintenance. In *Proc. of the International Conference on Automated Planning and Scheduling*.

Triantaphyllou, E. (2013). *Multi-criteria decision making methods: a comparative study*, volume 44. Springer Science & Business Media.

Volker, L., Altamirano, M., Herder, P., and van der Lei, T. (2014). The impact of innovative contracting on asset management of public infrastructure networks. In *Engineering Asset Management*, pages 665–676. Springer.